

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 20:56:55 ; Search time 66 Seconds  
(without alignments)  
4152.592 Million cell updates/sec

Title: US-10-026-021-2

Perfect score: 5078

Sequence: 1 MATCGEKREDFKVGNLGK.....KLQCLSSILMFSPNPNPH 970

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp19808:\*\n2: geneseqp19908:\*\n3: geneseqp20008:\*\n4: geneseqp20018:\*\n5: geneseqp20028:\*\n6: geneseqp20038:\*\n7: geneseqp20048:\*\n8: geneseqp20058:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5075	99.9	970	AAW78833	Aam78833 Human pro
2	5075	99.9	970	AAW78833	Aam78833 Human pro
3	5075	99.9	980	AAW79817	Aam79817 Human pro
4	5075	99.9	980	AAW41030	Aam41030 Human pro
5	3927.5	77.3	925	AAW92177	Aa92177 SAK-a ser
6	3927.5	77.3	925	ABBS7273	Abb57273 Mouse isc
7	1926	37.9	464	AAW92214	Aa92214 SAK-b ser
8	1883.5	37.1	416	AAW92176	Aa92176 SAK serin
9	1341.5	26.4	769	ABBS3067	Abb63067 Drosophi1
10	599.5	11.8	576	ABBS6250	Abb62520 Drosophi1
11	590.5	11.6	607	AAW34495	Aa34495 Human PRK
12	590.5	11.6	607	ADA50174	Ada50174 Human PRK
13	577	11.4	626	AAW35805	Aab35805 Protein f
14	560.5	11.0	685	AAW88432	Aaw88432 Disease a
15	560.5	11.0	685	AAW00915	Aay00915 Human bet
16	560.5	11.0	685	ABP61474	Abp61474 Human NF-
17	560.5	11.0	753	ABP41992	Abp41992 Human ova
18	547	10.8	603	AAW79317	Aau79317 Mouse pol
19	545	10.7	603	AAW79306	Aau79306 Mouse pol
20	545	10.7	603	AAW79318	Aau79318 Mouse pol
21	540	10.6	603	AAW79312	Aau79312 Mouse pol
22	540	10.6	603	AAW79320	Aau79320 Mouse pol
23	540	10.6	603	AAW79314	Aau79314 Mouse pol
24	540	10.6	603	AAW79310	Aau79310 Mouse pol
25	539	10.6	603	AAW79309	Aau79309 Mouse pol

26	539	10.6	603	5	AAU79316	Aau79316 Mouse pol
27	539	10.6	603	5	AAU79319	Aau79319 Mouse pol
28	539	10.6	603	5	AAU79308	Aau79308 Mouse pol
29	538	10.6	603	5	AAU79311	Aau79311 Mouse pol
30	537	10.6	603	5	AAU79315	Aau79315 Mouse pol
31	536	10.6	603	5	ABR48196	Abt48196 Human bla
32	533	10.5	603	5	AAU79313	Aau79313 Mouse pol
33	533	10.5	603	6	ABR92138	Abt92138 Human cer
34	528	10.4	603	2	AAW74620	Aaw74620 Human lun
35	526.5	10.4	603	2	ABBS3219	Abb33219 Protein s
36	524	10.3	523	6	AAW03425	Aab03425 Wheat put
37	524	10.3	523	6	ABR40719	Abt40719 Triticum
38	521.5	10.3	739	4	ABBS3486	Abb63486 Drosophi1
39	516.5	10.2	329	3	AAW56690	Aab56690 Human pro
40	516.5	10.2	832	4	AAW68340	Aab68340 A human d
41	516.5	10.2	832	5	AAW19053	Aae19053 Drosophi1
42	510.5	10.1	402	7	ABR61579	Abt61579 Human Hsa
43	510.5	10.1	402	7	ADD89969	Aad89969 Human can
44	510	10.0	1349	4	AAW70854	Aaw70854 C albican
45	508.5	10.0	403	2	AAW18084	Aaw18084 Human Aur

## ALIGNMENTS

RESULT 1	AAW78833	standard; protein; 970 AA.
ID	AAW78833	
XX	AAW78833;	
AC	AAW78833;	
XX		
DT	06-NOV-2001 (first entry)	
XX		
XX	Human protein SEQ ID NO 1495.	
DE		
XX		
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;	
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;	
KW	nervous system disorder; arthritis; inflammation.	
XX		
OS	Homo sapiens.	
XX		
PN	MO200157190-A2.	
XX		
PD		
XX	09-AUG-2001.	
XX		
PF	05-FEB-2001; 2001WO-US004098.	
XX		
PR	03-FEB-2000; 2000US-00496914.	
PR	27-APR-2000; 2000US-00560875.	
PR	20-JUN-2000; 2000US-00598075.	
PR	19-JUL-2000; 2000US-00620325.	
PR	01-SEP-2000; 2000US-00654936.	
PR	15-SEP-2000; 2000US-00663561.	
PR	20-OCT-2000; 2000US-00693325.	
PR	30-NOV-2000; 2000US-00728422.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y,	
PI	Ma Y, Zhao QA, Wang D, Wang J, Ren F, Chen R, Wang ZW,	
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R,	
XX	WPI, 2001-476283/51.	
DR	N-PSDB; AAK51966.	
XX		
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful	
PT	in diagnosis and gene therapy.	
XX		
PS	Claim 20, Page 3782-3784; 6221pp; English.	
XX		
CC	The invention relates to polynucleotides (AAK51456-AAK53435) and the	
CC	encoded polypeptides (AAW78323-AAW80302) that exhibit activity relating to	

CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polypeptides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
 CC (AAK52582) and 3666 (AAM00020) are omitted as the relevant pages from the  
 CC sequence listing were missing at the time of publication  
 CC  
 XX  
 SQ Sequence 970 AA;

Query Match 99.9%; Score 5075; DB 4; Length 970;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 969; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MATCGEKIEPKVGNLKGSGFAGVYRAESIHFGLEVALIMDKAMYKAGVQVQNE 60
DB 1 MATCGEKIEPKVGNLKGSGFAGVYRAESIHFGLEVALIMDKAMYKAGVQVQNE 60
QY 61 VKIHCOLGPSILLEYNFEDSNVYVLEMGCHNGEMNRVYKQVKEPSEARHFMQI 120
DB 61 VKIHCOLGPSILLEYNFEDSNVYVLEMGCHNGEMNRVYKQVKEPSEARHFMQI 120
QY 121 ITGMLYLSHSGILHRDITLSMLITRNMNIKIADFGLATQKMPHEKHVTLCTGPNYISP 180
DB 121 ITGMLYLSHSGILHRDITLSMLITRNMNIKIADFGLATQKMPHEKHVTLCTGPNYISP 180
QY 181 EIATRSAGHGESDVWSIGCMFYTLIGRPEDDTVYNTKVKVLDYEMPSLSIBAXD 240
DB 181 EIATRSAGHGESDVWSIGCMFYTLIGRPEDDTVYNTKVKVLDYEMPSLSIBAXD 240
QY 241 LIHQLRNPNADRLSSVLDHPMSNSSTKSDLGTVEDSIDSGHATISTATASST 300
DB 241 LIHQLRNPNADRLSSVLDHPMSNSSTKSDLGTVEDSIDSGHATISTATASST 300
QY 301 SISGSLFDRKRLIGQLEPKMTVPKNSSTDPSSSGGNSGYTQWQOETNSGGRV 360
DB 301 SISGSLFDRKRLIGQLEPKMTVPKNSSTDPSSSGGNSGYTQWQOETNSGGRV 360
QY 361 IQDAERPHSRVYRAYSDDSGTSNSQOAKTYMERGSAEMLSVSRSGSGGNEERY 420
DB 361 IQDAERPHSRVYRAYSDDSGTSNSQOAKTYMERGSAEMLSVSRSGSGGNEERY 420
QY 421 SPTDNNANFNPEKETSSSGSEFERPDNNQALSNHLCPGKTPFPADPTPQETVQWF 480
DB 421 SPTDNNANFNPEKETSSSGSEFERPDNNQALSNHLCPGKTPFPADPTPQETVQWF 480
QY 481 GNLQINAHLRKTTYDSISPNRDFQHPDLQKOTSKNMTDKYKKNSDASNAHSVQO 540
DB 481 GNLQINAHLRKTTYDSISPNRDFQHPDLQKOTSKNMTDKYKKNSDASNAHSVQO 540
QY 541 NTMKWTALHSPKPEIIQOECVFGSDPLSEOSKTRGMEPPWGYQNTLTSIPLVAHLK 600
DB 541 NTMKWTALHSPKPEIIQOECVFGSDPLSEOSKTRGMEPPWGYQNTLTSIPLVAHLK 600
QY 601 PIROKTKKAVVSIIDSEECVELVEKVASOEYVKEVLISSDGTTITTYNGRGFPLA 660
DB 601 PIROKTKKAVVSIIDSEECVELVEKVASOEYVKEVLISSDGTTITTYNGRGFPLA 660
QY 661 DRPSPPTNISRYSDNTPKTYWKRYQYASRFVOLVRSKPKITYFTFRYAKCIMENSPG 720
DB 661 DRPSPPTNISRYSDNTPKTYWKRYQYASRFVOLVRSKPKITYFTFRYAKCIMENSPG 720
QY 721 ADPEVWFYDGVKIHTEDFIOVIEKTKSYTLKSESVNSIKKEIKYMDHANEGHRL 780
DB 721 ADPEVWFYDGVKIHTEDFIOVIEKTKSYTLKSESVNSIKKEIKYMDHANEGHRL 780
QY 781 ALBESIISEERKTSAPFPPIIIGRKPGSTSPALSPPSVDSNVTTRDRASFNRMVH 840
DB 781 ALBESIISEERKTSAPFPPIIIGRKPGSTSPALSPPSVDSNVTTRDRASFNRMVH 840

```

```

QY 841 SAASPTOAPILNPSMTNBSGLTTTASGTDISSNSLKQCLPKSAQLKSVFKNVGMAT 900
DB 841 SAASPTOAPILNPSMTNBSGLTTTASGTDISSNSLKQCLPKSAQLKSVFKNVGMAT 900
QY 901 QLTSGAVWQVFDNGSQLVQAGVSSISYTPNGQTRRYGENEKLDPYIKOKLOCLSSITL 960
DB 901 QLTSGAVWQVFDNGSQLVQAGVSSISYTPNGQTRRYGENEKLDPYIKOKLOCLSSITL 960
QY 961 MFSNPINPH 970
DB 961 MFSNPINPH 970

```

RESULT 2  
 AAM39244  
 ID AAM39244 standard; protein; 970 AA.  
 XX  
 AC AAM39244;  
 XX  
 DT 22-OCT-2001 (first entry)  
 DE Human polypeptide SEQ ID NO 2389.  
 XX  
 DE Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW leukinaemia; thrombolytic; drug screening; arthritis; inflammation;  
 KW  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US034263.  
 XX  
 PR 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00489725.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-00598042.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00693036.  
 PR 29-NOV-2000; 2000US-00727344.  
 XX

(HYSB-) HYSBQ INC.

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,  
 Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
 Zhou P, Goodrich R, Drmanac RT;  
 WPI; 2001-442253/47.  
 DR N-PSDB; AAI58400.  
 XX

Novel nucleic acids and polypeptides, useful for treating disorders such  
 as central nervous system injuries.

Example 4; SEQ ID NO 2389; 10078pp; English.

The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
 encoded polypeptides (AAM38642-AAW42213) with nocotropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemia and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification

Query Match	Similarity	99.9%	Score	5075	DB	4	Length	970	
Best Local		99.9%	Pred.	No. 0					
Matches	969	Conservative	1	Mismatches	0	Indels	0	Gaps	0
Qy	1	MATCIGEKIEDPKYGNLIGKGSFAGVYRAESIHTGLAEVALIKMIDKKAMTKAGMYQVONE	60						
Db	1	MATCIGEKIEDPKYGNLIGKGSFAGVYRAESIHTGLAEVALIKMIDKKAMTKAGMYQVONE	60						
Qy	61	VKHICOLKHPBILLEYNYFEDSNVYVYLVLKCHNGEMNRILKNRKPPSENEARHFMQI	120						
Db	61	VKHICOLKHPBILLEYNYFEDSNVYVYLVLKCHNGEMNRILKNRKPPSENEARHFMQI	120						
Qy	121	ITGMVLYHSHOILHRDLTLSMLLTRNNNIKIADPGLATOLKMPPEKHYTLGTYNTYISP	180						
Db	121	ITGMVLYHSHOILHRDLTLSMLLTRNNNIKIADPGLATOLKMPPEKHYTLGTYNTYISP	180						
Qy	181	EIATRSAHGLESVDWLSGCMFYTLIGRPEDTDIVKNTLKNVVLADYEMSPFSLIEAKD	240						
Db	181	EIATRSAHGLESVDWLSGCMFYTLIGRPEDTDIVKNTLKNVVLADYEMSPFSLIEAKD	240						
Qy	241	LIHOLLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGATTSTALTAASST	300						
Db	241	LIHOLLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGATTSTALTAASST	300						
Qy	301	SISGSLFDPKRLLLIGQPLPNKQVYPPKKSSTDFSSSDGNSFTYQMGNOEFTNSGGRV	360						
Db	301	SISGSLFDPKRLLLIGQPLPNKQVYPPKKSSTDFSSSDGNSFTYQMGNOEFTNSGGRV	360						
Qy	361	IQDAEERPHSHYLRRAVSDRSSTNSGSAKTYTMEBCHSAEMLSVSKRSGGGENERY	420						
Db	361	IQDAEERPHSHYLRRAVSDRSSTNSGSAKTYTMEBCHSAEMLSVSKRSGGGENERY	420						
Qy	421	SPTNNNANIENFPEKETSSTSSGSPERRPNNNALSHLCPKGTTPFPADPTQTEVVOQWF	480						
Db	421	SPTNNNANIENFPEKETSSTSSGSPERRPNNNALSHLCPKGTTPFPADPTQTEVVOQWF	480						
Qy	481	GNLOINHLRKTTEXDSDSPMRDFGHPDLQDQTSKAMTPTKYKQNSDASDNASHVQO	540						
Db	481	GNLOINHLRKTTEXDSDSPMRDFGHPDLQDQTSKAMTPTKYKQNSDASDNASHVQO	540						
Qy	541	NTMYKMTALMSKPEIIQOECVFGSDPLSEQSKTRGMEBPWGQNTLTLSITSPVLAHLX	600						
Db	541	NTMYKMTALMSKPEIIQOECVFGSDPLSEQSKTRGMEBPWGQNTLTLSITSPVLAHLX	600						
Qy	601	PIROKTKKAVSIIIDSEVCVELYKVASOEVYKAVLOISDQNTITITTYPRNGRGFPLA	660						
Db	601	PIROKTKKAVSIIIDSEVCVELYKVASOEVYKAVLOISDQNTITITTYPRNGRGFPLA	660						
Qy	661	DRPSPDNDISRYSPDNLPEKYMRKYOYASRFVOLVRSPKPIITYFTYRKYACILMENSFG	720						
Db	661	DRPSPDNDISRYSPDNLPEKYMRKYOYASRFVOLVRSPKPIITYFTYRKYACILMENSFG	720						
Qy	721	ADFEWVFDGVKIHKTEDFIOVIEKTKGSYTLKSEBVSNLKEBIKMTMDHANEHRICTL	780						
Db	721	ADFEWVFDGVKIHKTEDFIOVIEKTKGSYTLKSEBVSNLKEBIKMTMDHANEHRICTL	780						
Qy	781	ALBSIIISEEKTKNSAPFPPIIIGKKPSTSPKLSPPPSVDSTNYPRDRASFRRMWH	840						
Db	781	ALBSIIISEEKTKNSAPFPPIIIGKKPSTSPKLSPPPSVDSTNYPRDRASFRRMWH	840						
Qy	841	SAASPQOPIINPMSVNTGGLATTTAGSTDISNSLDCLPKSAQLLKSYPVKNVGNAT	900						
Db	841	SAASPQOPIINPMSVNTGGLATTTAGSTDISNSLDCLPKSAQLLKSYPVKNVGNAT	900						
Qy	901	QUTSGAVVWVGQFNDSQLVVQGVSSISYTSPNGQTRRYGENEKLDPYIKQIKQLCLSSILL	960						

D <sub>b</sub>	901	QLTSGAVNQPFDSQLVVQACVSSITSTENGQTRRGNEKLPDYIKQKQCLSSILL	960
Q <sub>y</sub>	961	MPSNPPTNFH	970
D <sub>b</sub>	961	MPSNPPTNFH	970

Query Match	Best Local Similarity	99.9%	Score 5075	DB 4	Length 980
Matches 969	Conservative 1	Mismatches 0	Indels 0	Gaps 0	

```
QY 1 MATCIGEKIEDFKVGNLKGKSPAGVYRAESHTGLEVAIKMIDKAMTKAGVQVONE 60
Db 11 MATCIGEKIEDFKVGNLKGKSPAGVYRAESHTGLEVAIKMIDKAMTKAGVQVONE 70
QY 61 VKIHQOLKHPSELIELNYFEDSNVYVLYLEMCHNGEMNRYLKNRVKPFSENEARHPHQI 120
Db 71 VKIHQOLKHPSELIELNYFEDSNVYVLYLEMCHNGEMNRYLKNRVKPFSENEARHPHQI 130
QY 121 ITGMLYHSHGILHRDITLSNLLTRMNTKIDAPGLATQKMHGHTLCCGPNYISP 180
Db 131 ITGMLYHSHGILHRDITLSNLLTRMNTKIDAPGLATQKMHGHTLCCGPNYISP 190
QY 181 EIATRSAGHLESVDWLSGCMFYTLIGRPEDDTVKNATANKVYLADYEMPSFLSIAKD 240
Db 191 EIATRSAGHLESVDWLSGCMFYTLIGRPEDDTVKNATANKVYLADYEMPSFLSIAKD 250
QY 241 LIIHQLARNPADRLSSVLDHPMSNNSSTKSDLCCTVEDSIDSGHATTSTAITSAST 300
Db 251 LIIHQLARNPADRLSSVLDHPMSNNSSTKSDLCCTVEDSIDSGHATTSTAITSAST 310
QY 301 SISGLFDRKRLIIGQPLPNKMTVPFKNKSSTDPSSSGDENSFTOMGNETNSGRGV 360
Db 311 SISGLFDRKRLIIGQPLPNKMTVPFKNKSSTDPSSSGDENSFTOMGNETNSGRGV 370
QY 361 IQDAERPHSRYLRRAYSSDRSGTSSNSQOAKTYTWERCHSABMLSVSKRSGGENEBRY 420
Db 371 IQDAERPHSRYLRRAYSSDRSGTSSNSQOAKTYTWERCHSABMLSVSKRSGGENEBRY 430
QY 421 SPFDNNANINFPFKETSSSSSGFERPDNNQALSNHLCPKTKPPPADPPOETVQOMF 480
Db 431 SPFDNNANINFPFKETSSSSSGFERPDNNQALSNHLCPKTKPPPADPPOETVQOMF 490
QY 481 GNIQINAHLRKTEYDISISPNRDPQGHPLQKDTSKNMTDTVKKNKSDASDNASHVKQ 540
Db 491 GNIQINAHLRKTEYDISISPNRDPQGHPLQKDTSKNMTDTVKKNKSDASDNASHVKQ 550
QY 541 NTMKWTALHSKREIIQOECVRSDDLSEOSKTRGMEPPWGOQNTLRISITSPVVAHRLK 600
Db 551 NTMKWTALHSKREIIQOECVRSDDLSEOSKTRGMEPPWGOQNTLRISITSPVVAHRLK 610
QY 601 PIROKTKKAVSILDSEEVCEVELVKEVASOEYKVELOJISSQONTITTYPPNGGGFPLA 660
Db 611 PIROKTKKAVSILDSEEVCEVELVKEVASOEYKVELOJISSQONTITTYPPNGGGFPLA 670
QY 661 DRPSTPTDNISRYSPNLPKTKWRYQVARSFVQVLRSSPKITYFTRVAKCIIEMENSPG 720
Db 671 DRPSTPTDNISRYSPNLPKTKWRYQVARSFVQVLRSSPKITYFTRVAKCIIEMENSPG 730
QY 721 ADPEVWFYDGVKIHKTEDFIOVIEKTKGKYTLKSESVNSLKEEIKMYMDANEHRICL 780
Db 731 ADPEVWFYDGVKIHKTEDFIOVIEKTKGKYTLKSESVNSLKEEIKMYMDANEHRICL 790
QY 781 ALESIISEERKTRSAFPFPIIIGRKPSTSSPKALSPPSVDNSVYTRDRASFRMVMH 840
Db 791 ALESIISEERKTRSAFPFPIIIGRKPSTSSPKALSPPSVDNSVYTRDRASFRMVMH 850
QY 841 SAASPTQAPILNPSWMTNBEGLTTTASGTDISSNSLDCPLPKSAQILLKSVFKNVGMAT 900
Db 851 SAASPTQAPILNPSWMTNBEGLTTTASGTDISSNSLDCPLPKSAQILLKSVFKNVGMAT 910
QY 901 QUTSGAVWQFNDGSQLVQAGVSSISYTSPPNGOTTRVGENEKLPDYIKOYLQCCISLTL 960
Db 911 QUTSGAVWQFNDGSQLVQAGVSSISYTSPPNGOTTRVGENEKLPDYIKOYLQCCISLTL 970
QY 961 MFSNPTPNFH 970
Db 971 MFSNPTPNFH 980
```

```
RESULT 4
AA041030
ID AA041030 standard; protein; 980 AA.
```

```
XX AC AA041030;
XX XX 22-OCT-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 5961.
XX XX
XX KW Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX XX
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX XX
XX PF 26-DEC-2000; 2000MO-US034263.
XX XX
XX PR 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-00552317.
XX PR 20-JUN-2000; 2000US-00598042.
XX PR 19-JUL-2000; 2000US-00620312.
XX PR 03-AUG-2000; 2000US-00653450.
XX PR 14-SEP-2000; 2000US-00662191.
XX PR 19-OCT-2000; 2000US-00693036.
XX PR 29-NOV-2000; 2000US-00727344.
XX XX
XX PA (HYSE-) HYSEQ INC.
XX XX
XX PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Zhou J, Wang J, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Dirmacac RT;
XX XX
XX DR WPI; 2001-442253/47.
XX DR N-PSDB; AA160186.
XX XX
XX PT Novel nucleic acids and polypeptides, useful for treating disorders such
XX PT as central nervous system injuries.
XX XX
XX PS Example 2; SEQ ID NO 5961; 10078pp; English.
XX XX
XX CC The invention relates to human nucleic acids (AA157798-AA161369) and the
XX CC encoded polypeptides (AA038642-AA042213) with nocotropic,
XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide
XX CC of the invention may be used to treat diseases of the peripheral nervous
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: immune system suppression,
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC C.N.S disorders. Note: The sequence data for this patent did not form
XX CC part of the printed specification
XX XX
XX SQ Sequence 980 AA;
```

```
Query Match 99.9%; Score 5075; DB 4; Length 980;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 963; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATCIGEKIEDFKVGNLKGKSPAGVYRAESHTGLEVAIKMIDKAMTKAGVQVONE 60
Db 11 MATCIGEKIEDFKVGNLKGKSPAGVYRAESHTGLEVAIKMIDKAMTKAGVQVONE 70
QY 61 VKIHQOLKHPSELIELNYFEDSNVYVLYLEMCHNGEMNRYLKNRVKPFSENEARHPHQI 120
```



```

Db 71 VKIHQQLKHPSTILELVNYFEDSNVYVLVLEMGCHNGEMNRLLKRRVYKPFSENEARHPHQI 130
Qy 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATQKMPHEKHYTLCTPNYISP 180
Db 131 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATQKMPHEKHYTLCTPNYISP 190
Qy 181 EIAIRSAHGLSDVMSLGCMEFYTLIGRPFDITVNTLNTKVLADYEMPSFLSIAKD 240
Db 191 EIAIRSAHGLSDVMSLGCMEFYTLIGRPFDITVNTLNTKVLADYEMPSFLSIAKD 250
Qy 241 LTHQLLRNPADRLSLSSVLDHPFMSRNSSTSKDGLGYEDSIDSGHATISTAITASSST 300
Db 251 LTHQLLRNPADRLSLSSVLDHPFMSRNSSTSKDGLGYEDSIDSGHATISTAITASSST 310
Qy 301 SISGSLPDKRLLIGPLPKMTVPFKNKSSTDPSSSGDGNSTYTGQNGOETNSGGRV 360
Db 311 SISGSLPDKRLLIGPLPKMTVPFKNKSSTDPSSSGDGNSTYTGQNGOETNSGGRV 370
Qy 361 IDAERPSRRLRYRAYSDDSGTSSQOAKTYMERCHSAEMLSVSKRSGGGBERY 420
Db 371 IDAERPSRRLRYRAYSDDSGTSSQOAKTYMERCHSAEMLSVSKRSGGGBERY 430
Qy 421 SPTDNNANIFNFKETSSSSGSPERPDNNQALSNHLCPGKTEPPADPTPOTETVQOMF 480
Db 431 SPTDNNANIFNFKETSSSSGSPERPDNNQALSNHLCPGKTEPPADPTPOTETVQOMF 490
Qy 481 GNLQINAHARKTTEYDISPNRDFQGHPIQKDTSKNATDTYKKNQSDANASVKKQ 540
Db 491 GNLQINAHARKTTEYDISPNRDFQGHPIQKDTSKNATDTYKKNQSDANASVKKQ 550
Qy 541 NMTKWTALHSKEIIOQECVFGSDPLSEGSKTRGMEPPMGYONRLTASTSLVLAHRLK 600
Db 551 NMTKWTALHSKEIIOQECVFGSDPLSEGSKTRGMEPPMGYONRLTASTSLVLAHRLK 610
Qy 601 PIROKTKKAVNSILDESEVCEVLEKVEYASQEVYKEVLQISSDGNITITYYPNCGRGFPLA 660
Db 611 PIROKTKKAVNSILDESEVCEVLEKVEYASQEVYKEVLQISSDGNITITYYPNCGRGFPLA 670
Qy 661 DRPSPPTDNISRYSPDNLPKRYKQYASRFQVLRVSKPKITTYTRAKCILMENSFG 720
Db 671 DRPSPPTDNISRYSPDNLPKRYKQYASRFQVLRVSKPKITTYTRAKCILMENSFG 730
Qy 721 ADEEVPYGVKTHKEDFTQVLEKTKGSKYTLKSESEVNSLKEIKMYMDHANEGRICL 780
Db 731 ADEEVPYGVKTHKEDFTQVLEKTKGSKYTLKSESEVNSLKEIKMYMDHANEGRICL 790
Qy 781 ALESIISSEERKTRSAFPFIIIGRRKPGSTSPKALSPPPSVDSNYPTRDRASFNRMVMH 840
Db 791 ALESIISSEERKTRSAFPFIIIGRRKPGSTSPKALSPPPSVDSNYPTRDRASFNRMVMH 850
Qy 841 SAASPTQADILNBSMTNBSGLGTTTASGTDISSNSLKDCLPKSAOLKSVFKNVGMAT 900
Db 851 SAASPTQADILNBSMTNBSGLGTTTASGTDISSNSLKDCLPKSAOLKSVFKNVGMAT 910
Qy 901 QLTSGVWVQFNDGSLVQAGVSSISYSPNGQTRRYGENEGLPYIYKQKLOCLSSILL 960
Db 911 QLTSGVWVQFNDGSLVQAGVSSISYSPNGQTRRYGENEGLPYIYKQKLOCLSSILL 970
Qy 961 MFSNPTNPFH 970
Db 971 MFSNPTNPFH 980

```

RESULT 5  
 AAR92177  
 ID AAR92177 standard; protein; 925 AA.  
 AC AAR92177;  
 XX  
 DT 25-MAY-1996 (first entry)  
 XX  
 DE SAK-a serine-threonine kinase.

```

XX SAK-a; serine-threonine kinase; STK; agonist; antagonist;  

KM proliferative disease; cancer; tumour; antisense; transgenic animal;  

KW therapy.  

OS Mus musculus.  

XX  

FH Key Location/Qualifiers  

FT Domain 1..416  

FT /label=N-terminal domain  

FT /note="contains the kinase domain"  

FT Domain 417..925  

FT /label=C-terminal domain  

FT /note="contains 3 PEST regions"  

XX  

PN CA2150789-A.  

XX  

PD 03-DEC-1995.  

XX  

PE 01-JUN-1995; 95CA-02150789.  

XX  

PR 02-JUN-1994; 94US-00252995.  

XX  

PA (MOUN ) MOUNT SINAI HOSPITAL CORP.  

XX  

PI Dennis JM, Hefnerman M, Fode C;  

XX  

DR WPI, 1996-129817/14.  

DR N-PSDB; AAT08711.  

XX  

XX Nucleic acid encoding SAK serine-threonine kinase - useful for  

PT identifying modulators potentially useful in treatment or prevention of  

PT proliferative disease.  

PS Claim 5; Page 50-55; 73pp; English.  

XX  

XX 2 Isoforms, SAK-a (AAR92177) and SAK-b (AAR92214), of a novel  

CC serine/threonine kinase are associated with mitotic and meiotic cell  

CC division and are characterized by having a kinase domain at the N-  

CC terminus and 3 PEST sequences (rich in Pro, Ser, Thr, Asp, Glu) at the C-  

CC terminus. The N-terminal regions of the 2 isoforms are identical (see  

CC AAR92176). SAK-a and SAK-b can be obtained in recombinant form by expression  

CC of encoding sequences (see AAT08711-12), and used to test for inhibitory  

CC or stimulatory effects in the treatment/diagnosis of proliferative  

CC disorders such as cancer and viral (esp. HIV) infection, or used to  

CC raise antibodies  

XX  

SQ Sequence 925 AA;  

Query Match 77.3%; Score 3927.5; DB 2; Length 925;  

Best Local Similarity 78.6%; Pred. No. 1,1e-288;  

Matches 763; Conservative 76; Mismatches 83; Indels 49; Gaps 9;  

Qy 1 MATCIGEKIEDFVGNLTKGSGFAGVYRAESIHTGLEVAIKIMDKKAMTKAGVQVONE 60  

Db 1 MAACIGERIEDFVGNLTKGSGFAGVYRAESIHTGLEVAIKIMDKKAMTKAGVQVONE 60  

Qy 61 VKIHQQLKHPSTILELVNYFEDSNVYVLVLEMGCHNGEMNRLLKRRVYKPFSENEARHPHQI 120  

Db 61 VKIHQQLKHPSTILELVNYFEDSNVYVLVLEMGCHNGEMNRLLKRRVYKPFSENEARHPHQI 120  

Qy 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATQKMPHEKHYTLCTPNYISP 180  

Db 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATQKMPHEKHYTLCTPNYISP 180  

Qy 181 EIAIRSAHGLSDVMSLGCMEFYTLIGRPFDITVNTLNTKVLADYEMPSFLSIAKD 240  

Db 181 EIAIRSAHGLSDVMSLGCMEFYTLIGRPFDITVNTLNTKVLADYEMPSFLSIAKD 240  

Qy 241 LTHQLLRNPADRLSLSSVLDHPFMSRNSSTSKDGLGYEDSIDSGHATISTAITASSST 300  

Db 241 LTHQLLRNPADRLSLSSVLDHPFMSRNSSTSKDGLGYEDSIDSGHATISTAITASSST 300

```

```

QY 301 S I S G S L F D K R L L I G O P L P N K M T V F P K N K S T D F S S S G D G N S F Y T O W G N - - O E T S N S G R G 358
DB 301 S L S G S L I L D - R L L V G O P L P N K I T V F Q K N K S S D F - S S G D S S N F C T O W G N E Q A N S G R G 358
QY 359 R V I O D A E R P H S R L R A Y S D S G S N S O S O A K T T Y M E R C H S A E M L S V S K R S G G G E N E E 418
DB 359 R V I D A E R P H S R L R A H S S D R A S P S N - O S R A K T Y S V E R C H S Y E M L S K P R R S - - - - - 410
QY 419 R Y S P T D N N A N I F N F F E K E K T S S S G S F E R P D N N O A L S N H L C P G K T P P P A D P T P O T E T V Q 478
DB 411 - - - - - L D E N Q H S S N H H C L G K T P P P A D P T P O M E V Q 442
QY 479 W F G N L Q I N A H L R K T T E Y D S I S P N R D F Q G H P D L Q K O T S K N A M T D T K V K N S D A S D A N H S V K 538
DB 443 W F G N L Q I N A H L G E T N E H T V S P N R D F O D Y P D L Q - D T L R N A M T D T R A S K N A D T S A N V A Y K 501
QY 539 Q O N T M K M T A L H S K P E I I O E C V F G S D P L S E O S K T R G M E P P M G Y Q N R T L A S T S P L V A H R 598
DB 502 Q L S A M K T Y S A H H K P E V M Q E P - - G L H P H S B O S K N R S M E S T L G Y O K P T L R A S T S P L I A H R 559
QY 599 L K P I R O K T K A V A S I L D S E V C E L V E Y A S O E Y K A V E L Q I S S D G M T I T Y P N G R G P 658
DB 560 L K I R O K T K A V A S I L D S E V C E L L R E C S E G Y A K E L Q I S S D G M T I T Y P N D R G R P 619
QY 659 L A D R P S P T D N I S R Y S F D N L P E K Y M R K Y O Y A S R F O L V R S K S P K I T Y T F R Y A K C I I M E N S 718
DB 620 L A D R P L P T D N I S R Y S F D N L P E K Y M R K Y O Y A S R F I O L V A S K T P K I T Y T F R Y A K C I I M E N S 679
QY 719 P G A D F E V M Y P D G V K I H K T E D F I O V I E K T G S Y T L K E S E V N S I K E I K M Y M H A N G H R I 778
DB 680 P G A D F E V M Y P D G K I H K T E N L I H I E K T G I S Y N L K N E N E T S L K E E V K Y M H A N G H R I 739
QY 779 C L A L E S I I S E E R K T S A P F P I I I G R K G S T S S P A L S P P S V D S N Y P T R D A S F N R V 838
DB 740 C L S L E S I V E E E R K S G S S P F I I V G R K G N T S S P A L A P - V D S C C K G E R A S A S R S 798
QY 839 M S A S A P T O A P I L N P S M T N E G I G L T T T A S G D I S S N S L K D C L P K S A Q L L K S V F V N V G M 898
DB 799 V A S A A P T O S P G L S P T V E G I G H T A T A Y A G T G V S S - - - - - L P K S A Q L L K S V F V N V G M 853
QY 899 A T O L T S G A V W O P N D S O L V O A G V S I S Y T S P N G O T T R Y G E N E K L P D Y I K O L Q L S I 958
DB 854 A T O L T S G A V W O P N D S O L V O A G V S I S Y T S P D G O T T R Y G E N E K L P E Y I K O L Q L S I 913
QY 959 L I M F S N P T N F 969
DB 914 L I M F S N P T N F 924

```

## RESULT 6

```

ABBS7273
ID ABB57273 standard; protein; 925 AA.
AC ABB57273;
XX
XX 07-MAR-2002 (first entry)
DT
DE Mouse ischaemic condition related protein sequence SEQ ID NO:766.
XX
XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
XX vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
XX Mus musculus.
OS
XX
XX WO200188188-A2.
XX
XX 22-NOV-2001.
XX
XX 18-MAY-2001; 2001WO-JP004192.
XX
XX 18-MAY-2000; 2000JP-00145977.
XX
XX (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
PA

```

```

XX
PI Ishikawa K, Arai S, Takahashi Y, Nagata T, Ishii Y;
XX
XX WPI; 2002-034733/04.
DR N-PSDB; AB199713.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.
XX
PS Claim 2; Page 1893-1897; 2690pp; English.
XX
XX The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (1) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (1). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (AB199202 to AB199912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. AB199913 and AB199914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX
XX Sequence 925 AA;
XX
Query Match 77.3%; Score 3927.5; DB 5; Length 925;
Best Local Similarity 78.6%; Pred. No. 1.1e-288;
Matches 763; Conservative 76; Mismatches 83; Indels 49; Gaps 9;
QY 1 MATCIGEKIEDPKGNLIGKSPAGVYRAESIHTGLEVAIYKIDKKAYKAGVQVONE 60
DB 1 MAACIGERIEDFKGNLIGKSPAGVYRAESIHTGLEVAIYKIDKKAYKAGVQVONE 60
QY 61 VKIHQQLHPSTLIELYNNFEDSNVYLVLEMCHGEMNRYLKNVYKPSSEBAHFHQI 120
DB 61 VKIHQQLHPSTLIELYNNFEDSNVYLVLEMCHGEMNRYLKNVYKPSSEBAHFHQI 120
QY 121 ITGMVLHSHGILHRDLTLSNMLLTRNNIKIADFGLATOLKMPHEKHYYTLCGTPNYISP 180
DB 121 ITGMVLHSHGILHRDLTLSNMLLTRNNIKIADFGLATOLKMPHEKHYYTLCGTPNYISP 180
QY 181 EIATRSAGLESVDWGLCMFYTLILGPPPDYTVKNTLKVVLADYEMPSLIEAKD 240
DB 181 EIATRSAGLESVDWGLCMFYTLILGPPPDYTVKNTLKVVLADYEMPSLIEAKD 240
QY 241 LIHOLLRNPPADRLSLSSVLDPHPPMSRNSSTKSKDLGTVEDSINDGHATTSTATTASSST 300
DB 241 LIHOLLRNPPADRLSLSSVLDPHPPMSRNPSPKSKOVGVEDSMDSGHATLSTTTTASSGT 300
QY 301 S I S G S L F D K R L L I G O P L P N K M T V F P K N K S T D F S S S G D G N S F Y T O W G N - - O E T S N S G R G 358
DB 301 S L S G S L I L D - R L L V G O P L P N K I T V F Q K N K S S D F - S S G D S S N F C T O W G N E Q A N S G R G 358
QY 359 R V I O D A E R P H S R L R A Y S D S G S N S O S O A K T T Y M E R C H S A E M L S V S K R S G G G E N E E 418
DB 359 R V I D A E R P H S R L R A H S S D R A S P S N - O S R A K T Y S V E R C H S Y E M L S K P R R S - - - - - 410
QY 419 R Y S P T D N N A N I F N F F E K E K T S S S G S F E R P D N N O A L S N H L C P G K T P P P A D P T P O T E T V Q 478
DB 411 - - - - - L D E N Q H S S N H H C L G K T P P P A D P T P O M E V Q 442
QY 479 W F G N L Q I N A H L R K T T E Y D S I S P N R D F Q G H P D L Q K O T S K N A M T D T K V K N S D A S D A N H S V K 538
DB 443 W F G N L Q I N A H L G E T N E H T V S P N R D F O D Y P D L Q - D T L R N A M T D T R A S K N A D T S A N V A Y K 501
QY 539 Q O N T M K M T A L H S K P E I I O E C V F G S D P L S E O S K T R G M E P P M G Y Q N R T L A S T S P L V A H R 598
DB 502 Q L S A M K T Y S A H H K P E V M Q E P - - G L H P H S B O S K N R S M E S T L G Y O K P T L R A S T S P L I A H R 559

```

QY 599 LKPIROKTKKAVVSIIDSEVCEVLKVEYASQRYKVEVLQISSDGMTITTYENGGRFP 658  
 DB 560 LKPIROKTKKAVVSIIDSEVCEVLKVEYASQRYKVEVLQISSDGMTITTYENGGRFP 619  
 QY 659 LADRPSPPTDNISRYSPFNDLPEKYMWRKYQYASRFVLVNSKSPKITTYPTRYAKCIMENS 718  
 DB 620 LADRPSPPTDNISRYSPFNDLPEKYMWRKYQYASRFVLVNSKSPKITTYPTRYAKCIMENS 679  
 QY 719 PGADPFVWYDGVKIKHTEDFIOVIERKTSYTLKSESEVNSLKEEIKMYMDHANGHRI 778  
 DB 680 PGADPFVWYDGVKIKHTEDFIOVIERKTSYTLKSESEVNSLKEEIKMYMDHANGHRI 739  
 QY 779 CIALSEIIEEEKTSAPPEPPIIIRKQSGTSSPALSPPEVDNPNTRDASPNRV 838  
 DB 740 CIALSEIIEEEKTSAPPEPPIIIRKQSGTSSPALSPPEVDNPNTRDASPNRV 798  
 QY 839 MHSASPTAPILNPMVTNEGILGTTTASGTDISSNSLDCPKSAQLKSVFVKNVGM 898  
 DB 799 VNSAAPTOSPGSPSTVTEGIGHNTATGTCVSSS-----LPKSAQLKSVFVKNVGM 853  
 QY 899 ATQLTSGAVVQFNDGSQLVQAGVSSISYTSFNGQTRRYGENEKLPTYIKQKLOCLSSI 958  
 DB 854 ATQLTSGAVVQFNDGSQLVQAGVSSISYTSFNGQTRRYGENEKLPTYIKQKLOCLSSI 913  
 QY 959 LMFSPNTPNF 969  
 DB 914 LMFSPNTPNF 924

## RESULT 7

AA92214  
 ID AA92214 standard; protein; 464 AA.

AA92214;

25-MAY-1996 (first entry)

SAK-b serine-threonine kinase.

SAK-b; serine-threonine kinase; STK; agonist; antagonist;  
 proliferative disease; cancer; tumour; antisense; transgenic animal;  
 therapy.

Mus musculus.

Key Location/Qualifiers

FT 1..416 /label= "N-terminal domain"

FT /note= "contains the kinase domain"

FT 417..464 /label= "C-terminal domain"

FT /note= "contains 3 PEST regions"

CA2150789-A.

03-DEC-1995.

01-JUN-1995; 95CA-02150789.

02-JUN-1994; 94US-00252995.

(MOUN ) MOUNT SINAI HOSPITAL CORP.

PI Dennis JW, Heffernan M, Fode C;

DR WPI; 1996-129817/14.

XX N-PSDB; AAT08712.

PT Nucleic acid encoding Sak serine-threonine kinase - useful for  
 PT identifying modulators potentially useful in treatment or prevention of  
 PT proliferative disease.  
 XX

PS Claim 6; Page 58-61; 73pp; English.

CC 2 Isoforms, SAK-a (AA92177) and SAK-b (AA92214), of a novel  
 CC serine/threonine kinase are associated with mitotic and meiotic cell  
 CC division and are characterized by having a kinase domain at the N-  
 CC terminus and 3 PEST sequences (rich in Pro, Ser, Thr, Asp, Glu) at the C-  
 CC terminus. The N-terminal regions of the 2 isoforms are identical (see  
 CC AA92176). SAK-a and SAK-b can be covalently modified in recombinant form by expression  
 CC of encoding sequences (see AAT08712-12), and used to test for inhibitory  
 CC or stimulatory effects in the treatment, diagnosis of proliferative  
 CC disorders such as cancer and viral (esp. HIV) infection, or used to  
 CC raise antibodies

CC Sequence 464 AA;

QY Query Match 37.9%; Score 1926; DB 2; Length 464;  
 DB Best Local Similarity 84.9%; Pred. No. 3,3e-137;  
 DB Matches 372; Conservative 30; Mismatches 30; Indels 6; Gaps 5;

QY 1 MATCTGKTEDPRKGNLKGSPAGYRRESIHTGLEVAIKMIDKKMYKAGVORVOR 60

DB 1 MACIGERLEDPRVGNLKGSPAGYRRESIHTGLEVAIKMIDKKMYKAGVORVOR 60

QY 61 VKIHCOLKPSILDELNYFEDSNVYLVLEMCHNGEMNRYLNKRVKPFSENEARHFMQI 120

DB 61 VKIHCOLKPSVLELYNFEDNNVYLVLEMCHNGEMNRYLNKRVKPFSENEARHFMQI 120

QY 121 ITGMLYLHSHGILHRDLTSLNLLTRNNIKIADFGLATQLMKPEKHYTLGTPYVISP 180

DB 121 ITGMLYLHSHGILHRDLTSLNLLTRNNIKIADFGLATQLMKPEKHYTLGTPYVISP 180

QY 181 EIATRSAGHLSDDWSIGCMFYTLIGRPPTDPTKNTLVADYMPSPFLSIEARD 240

DB 181 EIATRSAGHLSDDWSIGCMFYTLIGRPPTDPTKNTLVADYMPSPFLSIEARD 240

QY 241 LIHQLLRNPADRLSLSVLDHPFMSRNSSTKSLGTYEDSIDSGHATITATSSST 300

DB 241 LIHQLLRNPADRLSLSVLDHPFMSRNSSTKSLGTYEDSIDSGHATITATSSST 300

QY 301 SIIGSLFDKRLILGQPLPKNTVFPKSKSTDFSSGCGNSFYTQGN--QETNSGRG 358

DB 301 SIIGSLFD--RRLLVGQPLPKNTVFPKSKSTDFSSGCGNSFYTQGN--QETNSGRG 358

QY 359 RVIQDAEERPHSRYLRRASPSDRSGTNSQSAKYTYMERCHSAEMLSYKSGGSENE 418

DB 359 RVIQDAEERPHSRYLRRASPSDRSGTNSQSAKYTYMERCHSAEMLSYKSGGSENE 418

QY 419 RYSPTDNNANIFNFEK 436

DB 417 RYSPTKSNVNVLTSLNTK 434

## RESULT 8

AA92176  
 ID AA92176 standard; protein; 416 AA.

AA92176;

25-MAY-1996 (first entry)

Sak serine-threonine kinase N-terminus.

Sak; serine-threonine kinase; STK; agonist; antagonist;  
 proliferative disease; cancer; tumour; antisense; transgenic animal;  
 therapy.

XX Mus musculus.

OS CA2150789-A.

XX 03-DEC-1995.

XX 01-JUN-1995; 95CA-02150789.

```

XX 02-JUN-1994; 94US-00252995.
XX (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX
XX Dennis JW, Heffernan M, Fode C;
XX WPI; 1996-129817/14.
XX N-PSDB; AAT08710.
XX
XX Nucleic acid encoding Sak serine-threonine kinase - useful for
XX identifying modulators potentially useful in treatment or prevention of
XX proliferative disease.
XX
XX Claim 3; Page 46-48; 73pp; English.
XX
XX 2 Isoforms, sak-a and sak-b, of a novel serine/threonine kinase have an
XX identical N-terminal sequence (AA92176) that contains the kinase domain
XX and that shows significant homology to the polo subfamily. The C-terminal
XX sequences (each contg. 3 PST regions) of the 2 isoforms differ (see
XX AA92177 and AA92214). Sak-a and Sak-b are associated with mitotic and
XX meiotic cell division, and may be involved in cell proliferation. They
XX can be obt'd. in recombinant form by expression of encoding sequences (see
XX AAT08710-12) and used to test for inhibitory or stimulatory cpds. that
XX may be useful in the treatment/diagnosis of proliferative disorders, such
XX as cancer and viral (esp. HIV) infection
XX
XX Sequence 416 AA;
XX
XX Query Match 37.1%; Score 1883.5; DB 2; Length 416;
XX Best Local Similarity 87.7%; Pred. No. 4.6e-134;
XX Matches 367; Conservative 27; Mismatches 19; Indels 5; Gaps 4;
XX
XX 1 MATCIGEKIDPFKVGNLGKGFAGVYRAESHTGLAEVAIKMIDKKAMTKAGVORVONE 60
XX 1 MAACIGERIEDPFKVGNLGKGFAGVYRAESHTGLAEVAIKMIDKKAMTKAGVORVONE 60
XX
XX 61 VKIHCOLKHPSTLELYNYPEDSNVYLVLEMCNGEMNRYLKNRYKVPSENEARHFMQI 120
XX 61 VKIHCOLKHPSTLELYNYPEDSNVYLVLEMCNGEMNRYLKNRYKVPSENEARHFMQI 120
XX
XX 61 VKIHCOLKHPSTLELYNYPEDSNVYLVLEMCNGEMNRYLKNRYKVPSENEARHFMQI 120
XX 61 VKIHCOLKHPSTLELYNYPEDSNVYLVLEMCNGEMNRYLKNRYKVPSENEARHFMQI 120
XX
XX 121 ITGMLYHSHGILHRDLTSLNLLTRNNIKIADFGLATQUMPHKHYTLGTPNYISP 180
XX 121 ITGMLYHSHGILHRDLTSLNLLTRNNIKIADFGLATQUMPHKHYTLGTPNYISP 180
XX
XX 121 ITGMLYHSHGILHRDLTSLNLLTRNNIKIADFGLATQUMPHKHYTLGTPNYISP 180
XX 121 ITGMLYHSHGILHRDLTSLNLLTRNNIKIADFGLATQUMPHKHYTLGTPNYISP 180
XX
XX 181 EATRSAGHLESVWSLGCFFITLLIGRPEDTDVKNLTNRYLADYEMPSFSLSEAKD 240
XX 181 EATRSAGHLESVWSLGCFFITLLIGRPEDTDVKNLTNRYLADYEMPSFSLSEAKD 240
XX
XX 181 EATRSAGHLESVWSLGCFFITLLIGRPEDTDVKNLTNRYLADYEMPSFSLSEAKD 240
XX 181 EATRSAGHLESVWSLGCFFITLLIGRPEDTDVKNLTNRYLADYEMPSFSLSEAKD 240
XX
XX 241 LTHQLRRPADRLSLSSVLDHPMSRNSSTKSLGTVEDSIDGHAITSTAITASSST 300
XX 241 LTHQLRRPADRLSLSSVLDHPMSRNSSTKSLGTVEDSIDGHAITSTAITASSST 300
XX
XX 241 LTHQLRRPADRLSLSSVLDHPMSRNSSTKSLGTVEDSIDGHAITSTAITASSST 300
XX 241 LTHQLRRPADRLSLSSVLDHPMSRNSSTKSLGTVEDSIDGHAITSTAITASSST 300
XX
XX 301 STSGSLFDRKRLILGQPLPNKMTVPKNSSTDPSGSGDGSFYOMGN--QETSNSGRG 358
XX 301 STSGSLFDRKRLILGQPLPNKMTVPKNSSTDPSGSGDGSFYOMGN--QETSNSGRG 358
XX
XX 301 STSGSLFDRKRLILGQPLPNKMTVPKNSSTDPSGSGDGSFYOMGN--QETSNSGRG 358
XX 301 STSGSLFDRKRLILGQPLPNKMTVPKNSSTDPSGSGDGSFYOMGN--QETSNSGRG 358
XX
XX 359 RYIDAEERPHSRYLRRAYSDRSGTSNSQAKTYTMRCHSAEMLSVKRSRG 411
XX 359 RYIDAEERPHSRYLRRAYSDRSGTSNSQAKTYTMRCHSAEMLSVKRSRG 411
XX
XX 359 RYIDAEERPHSRYLRRAYSDRSGTSNSQAKTYTMRCHSAEMLSVKRSRG 411
XX 359 RYIDAEERPHSRYLRRAYSDRSGTSNSQAKTYTMRCHSAEMLSVKRSRG 411
XX
XX RESULT 9
XX ABB63067
XX ID ABB63067 standard; protein; 769 AA.
XX
XX AC ABB63067;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 15993.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;

```

```

KW pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (BEK ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX
XX N-PSDB; ABL07170.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
XX Disclosure; SEQ ID NO 15993; 21pp + Sequence listing; English.
XX
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutic and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL20511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-
XX AB82202). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 769 AA;
XX
XX Query Match 26.4%; Score 1341.5; DB 4; Length 769;
XX Best Local Similarity 33.0%; Pred. No. 1.7e-92;
XX Matches 321; Conservative 156; Mismatches 234; Indels 261; Gaps 24;
XX
XX 6 GEKIDEPKVGNLGKGFAGVYRAESHTGLAEVAIKMIDKKAMTKAGVORVONEKIKC 65
XX 8 GETIEDIEVGHLLGKGFATVYKARCLTHQDVAKMIDKLIQGTGLNRYRQVEIHS 67
XX
XX 6 QLKHPSTLELYNYPEDSNVYLVLEMCNGEMNRYLKNRYKVPSENEARHFMQIITGM 125
XX 68 RLKHPSTVLYQTYFPODANYVYLVLELHNGELHRVWNIARPTETELASILKQVAGIL 127
XX
XX 126 YLHSHGILHRDLTSLNLLTRNNIKIADFGLATQUMPHKHYTLGTPNYISPRIATR 185
XX 128 YLHSHGILHRDLTSLNLLTRNNIKIADFGLATQUMPHKHYTLGTPNYISPRIATR 187
XX
XX 186 SAHGLSDVWSLGCFFITLLIGRPEDTDVKNLTNRYLADYEMPSFSLSEAKD 245
XX 188 TSHGLPADVWSVQMLYTLVGPRPETDAVOSTLKNVMSYINPAHLSYADQLINKL 247
XX
XX 246 LRANPADRLSLSSVLDHPMSRNSSTKSLGTVEDSIDGHAITSTAITASSSTISGS 305
XX 248 LKQDLPHERITLLEVLCHPMLKCS-----NGHSA----- 277
XX
XX 306 LFDKRLILGQPLPNKMTVPKNSSTD-----FSSGSGDGSFYOMGNQETSNSGRGRV 360
XX 278 -----PALNVFSQSMESGSGIITTFASSDRNSQOI-----RSVENSQPOOV 320
XX
XX 361 IODAEERPHSRYLRRAYSDRSGTSNSQAKTYTMRCHSAEMLSVKRSRGGENEBRY 420
XX 321 LPQIREEPKOVHKKLPY----- 337
XX
XX 421 SPTDNNANIFNFFKKTSSSGSFERRDNNALSNHLCPKTPPPPTPTOTETVOQWF 480

```

Db 338 -----BOTGLF-----GQASTGLAEP-----NWP 356  
 Qy 481 GNLIQNAHLKRTTEVDSISPNRDFQGHPLQKOTSKNMTDTVKKNASDASNAHSVQKQ 540  
 Db 357 GNAKSAFAMEA----- 368  
 Qy 541 NTMKYMTALHSHKREIIQOECVGSDELSEQSKTRGMEPPWGYONRTLRSTSPVLAHRLK 600  
 Db 369 -----GNVPSNQASL-----KEDRLSVPLANKRL 395  
 Qy 601 PIRQTKKAVNSIILDSSEVCELVK--EYASQETVKEVLIQISDGNITITITTYNGRGFP 658  
 Db 396 PRRYTKKAIMSILRNGEVLLEFKRPPTYNEDRINDICRISDDGQRIITIIYOPDPGRGLP 455  
 Qy 659 LADRPSS---PTDNISRYSEFDNLPEKYWRKYQYASRFVQLVRSKPKITPTFYACIIM 715  
 Db 456 VAEQPPDLQIPSGDCV-YNYDNLPSKMKKTYIGARFVGLVSKITKVTYFTLGKQIM 514  
 Qy 716 ENSPGADPEVWFYDGVKIKHT-EDPIQVIEKTKSYTLKSESEVNSLKEEIKYMDHANE 774  
 Db 515 EFM--TDFEIRFYSGAKLKTPEGLKAYDRNGMLLDVYSCSESRL-----IEHGR 565  
 Qy 775 GHRICLAEISISEERKTRSAFPPIITIGRKPSTSPKALSPPPSVDSNPTFRDPAF 834  
 Db 566 CFTHCVNINALEVAGTKNSC--FPVITIGRPIITD-----VQPAQLDG--LRDTTN- 614  
 Qy 835 NNMWHSASPTQAPILNBSMTNBEGLGTTTASGTDISSNLKDLPLPSAQ--LKSVF 892  
 Db 615 ---IAFSTPKSNQGS-NFSLST-----ISSTNTSDPGTNCGRSMLAHQNIPIKRI 665  
 Qy 893 VKNVGAATOLTSQAVWQFNDGSQLVV---QAGVSSISYTSNPGQTFYGENEKLPDYI 948  
 Db 666 VEIGIATLHSHGVVQVGYDSVSVISMOG--GITTYQNGSTHFGKGDLDLPFVY 723  
 Qy 949 KQKLOGLSSILL 960  
 Db 724 RDRVGQIPNIQL 735

## RESULT 10

ABB62520

ID ABB62520 standard; protein; 576 AA.

XX AC ABB62520;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polyprotein seq ID NO 14352.

XX KM Drosophila: developmental biology; cell signalling; insecticide;

XX KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE ) PE CORP NY.

XX PI Venter JC, Adams M, Li PMD, Myers EW;

XX DR WPI: 2001-656860/75.

XX DR N-PSDB; ABL06623.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.

XX PS Disclosure; SEQ ID NO 14352; 21np + Sequence Listing; English.  
 XX CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB16311), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 576 AA;

Qy Query Match 11.8%; Score 599.5; DB 4; Length 576;  
 Best Local Similarity 34.1%; Pred. No. 2.3e-36;  
 Matches 136; Conservative 71; Mismatches 121; Indels 71; Gaps 9;  
 Qy 2 ATCGEKIED-----FKVGNLLGKGSFAGVYRAESIHGTLEVAIKMIDKAMYKAGVQ 55  
 Db 9 STDIPLRLVDINQRTYKMRFFGKGFACKEIIDVEITDVPAGKIVSKLMIKINQKE 68  
 Qy 56 RVQNEVKIKCOLKPSILELYNYFEDSNVYVLEWCHNGENMYLKNVKEPSENEARH 115  
 Db 69 KTAQETTHRSLNHPVIVKFNHYFEDSQNIYIVLELCKRSWMLHKRR-KSITEPCRY 127  
 Qy 116 FNGQITGMLYHSHGILHRLDLSNLLTRMMNLIKADFGIATOLKMHKHYTLGCTP 175  
 Db 128 YVYQIIGVYKYLHDKRIIHRDLKGNLFNLDLHVKIGDFGLATRIEYGERKTKLCGP 187  
 Qy 176 NYISPEIATRSAGLESVDYSLGMEFYTLIGRPEDTDTVKNVTLNKVLAIDYEMSPFS 235  
 Db 188 NYIAPEILTKGHSFEVDIWSIGCVWYTLVGQPPETIKTLQDTSKIKCEKRVSYLR 247  
 Qy 236 IEAKDLIHQLRRNPADRLSLSSVLDHPFMS-----RNSSTKSDLG--TYEDSI 283  
 Db 248 KPAAADVIMLQPNPESRAIQGLNFEPLKSGSKVWPLPSQLTVAAPIGSNDTIEDSM 307  
 Qy 284 D-----SG-----HATISATITASSSTSISSGLFDKRLILIGOLPLPNKMT 323  
 Db 308 HRRKPLMNNGIRPDDTRLSTFLKANLHDAITASAVC----- 345  
 Qy 324 VEPKNSSTPSSSGDGNSEFYQMGNGQTSNNGRGVIG 362  
 Db 346 ----RHSEDYRS--DIESLYOQLTN--LINGKPRILQ 374

## RESULT 11

ID AAE34495 standard; protein; 607 AA.

XX AC AAE34495;

XX DT 14-MAY-2003 (first entry)

XX DE Human PRK protein.

XX KM Human; pyridylpyrimidine derivative; cellular protein kinase; Scrapie;  
 KM cellular protein phosphatase; cellular signal transduction; prophylaxis;  
 KM prion infection; chronic wasting disease; CWD; Creutzfeldt-Jacob disease;  
 KM CJD; transmissible mink encephalopathy; bovine spongiform encephalopathy;  
 KM TSE; BSE; Gerstmann-Strausler-Scheiner syndrome; GSS; Alpers syndrome;  
 KM fatal familial insomnia; FFI; kuru; neurodegenerative disease; nootropic;  
 KM Alzheimer's disease; PRK.

XX OS Homo sapiens.

XX PN WO200293164-A2.

XX PD 21-NOV-2002.





KW Grave's disease; lupus erythematosus; multiple sclerosis;  
 KW myasthenia gravis; osteoarthritis; osteoporosis; pancreatitis;  
 KW polycystic kidney disease; polymyositis; scleroderma;  
 KW Sjogren's syndrome; autoimmune thyroiditis; cancer; infection; trauma;  
 KW cell proliferation.

OS Homo sapiens.

XX Key Location/Qualifiers  
 FH Binding-site 89..96 /note="potential ATP binding site"  
 FT Peptide 506..511 /note="presumed regulatory sequence common to polo  
 FT family protein kinases"

XX WO9858052-A2.  
 XX 23-DEC-1998.  
 XX 19-JUN-1998; 98WO-US012813.  
 XX 19-JUN-1997; 97US-00878989.  
 XX (INCYTE) INCYTE PHARM INC.  
 XX Bandman O, Hillman JL, Corley NC, Guegler KJ, Lal P, Goli SK,  
 XX Shah P;  
 XX MPI; 1999-080952/07.  
 XX N-PSDB; AAX06831.

PT New disease associated protein kinases - used to stimulate cell  
 PT proliferation and to treat the immune response and cancer.

XX Claim 1; Page 54-56; 93pp; English.

PS This is the amino acid sequence of human disease associated protein  
 CC kinase DAPK-1, as deduced from a consensus sequence (see AAX06831) of  
 CC overlapping cDNA clones from libraries which are immortalised or  
 CC cancerous and show inflammatory or immune responses. DAPK-1 shows 53%  
 CC homology to human proliferation-related protein kinase PRK (GI 1488263).  
 CC The invention provides DAPK-1 to DAPK-7 polypeptides (see AAW8432-38)  
 CC and cDNA clones encoding them (see AAX06831-36 and AAX06882), as well as  
 CC expression vectors, host cells, agonists, antagonists and antibodies. The  
 CC invention further provides uses of such products in the diagnosis,  
 CC prevention and treatment of diseases associated with cell proliferation,  
 CC especially cancer or an immune response (claimed). Conditions that may be  
 CC treated include adult respiratory distress syndrome, allergies, asthma,  
 CC arteriosclerosis, bronchitis, emphysema, hyperesoinphilia, myocardial or  
 CC pericardial inflammation, rheumatoid arthritis, Addison's disease, AIDS,  
 CC anaemia, atherosclerosis, various diseases of the digestive system,  
 CC atopic dermatitis, dermatomyositis, diabetes mellitus,  
 CC glomerulonephritis, gout, Grave's disease, lupus erythematosus, multiple  
 CC sclerosis, myasthenia gravis, osteoarthritis, osteoporosis, pancreatitis,  
 CC polycystic kidney disease, polymyositis, scleroderma, Sjogren's  
 CC syndrome, autoimmune thyroiditis, complications of cancer, extracorporeal  
 CC circulation, viral, bacterial, fungal, parasitic, protozoal and  
 CC helminthic infections, and trauma (disclosed)

XX Sequence 685 AA;

Query Match 11.0%; Score 560.5; DB 2; Length 685;  
 Best Local Similarity 41.6%; Pred. No. 2.8e-33;  
 Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

QY 15 GNLIGKGSFAGVYRAESIHITGLVAIKIMDKKAMYKAGMORVONVYKIHQCLKHPSILE 74  
 DB 85 GKVIGKGGFACQCYEMTIDLTNNKVVAAKIIIPHSRAKCPQREKIDKEITELRHIIHKKHVQ 144  
 QY 75 LYNTFPEBSNYVYVLEMGCHNEENRKYLNKRVYPSENEBAHFHMQITGMYYLHSHGILH 134  
 DB 145 FYHYFEDENKITYLLECYGSRSSMAHILKAR-KVLTPEVRVRYLMDIVSGKLTHLEQELIH 203

QY 135 RDLTISNLLITRRNNIKIADFGLATQKMPHEKHYYLTCGTPNTYISPEIATRSAGHLESVD 194  
 DB 204 RDLTGKGFPEINAEMLKVGFGAARLEPLEHRRRTICGTPNYLSPEVLNKGHCESDI 263  
 QY 195 WSLCGMYTLLIGRPPTDVTIVKTLANKVVLADYEMSPSLSEKOLIHOLLRRNPADRL 254  
 DB 264 WALGCVMTMLGRPPETTNLKETRYCIRREARYTSPSLAPAKHLLASLNSKNPDRP 323

QY 255 SLSSVLDPHF 264  
 DB 324 SLDDIIRHDF 333

RESULT 15  
 AAY00915  
 ID AAY00915 standard; protein; 685 AA.  
 XX AAY00915;  
 XX 28-MAY-1999 (first entry)  
 XX Human serum inducible kinase.  
 XX Serum inducible kinase; SNK protein; human; proliferative disease;  
 KW leukaemia; solid tumour cancer; metastasis; chronic inflammatory;  
 KW psoriasis; rheumatoid arthritis; proliferative cardiovascular disease;  
 KW restenosis; ocular disorder; diabetic retinopathy; haemangioma;  
 KW benign hyperproliferative disease; diagnosis.

XX Homo sapiens.  
 OS WO9909146-A1.  
 XX 25-FEB-1999.  
 XX 20-AUG-1998; 98WO-US017248.  
 XX 20-AUG-1997; 97US-0056112P.  
 XX (SMK ) SMITHKLINE BEECHAM CORP.  
 XX Anderson KM, Jackson JR, Hansbury MJ, Nerurkar SS, Roshak AK,  
 XX Bouzyk M;  
 XX MPI; 1999-181027/15.  
 XX N-PSDB; AAX27227.

DR New serum inducible kinase (SNK) polypeptides and polynucleotides -  
 PT useful for treating proliferative diseases.

XX Claim 1; Page 39-40; 41pp; English.

PS This sequence is a human serum inducible kinase (SNK) of the invention.  
 CC The invention relates to diagnostic assays or kits for detecting diseases  
 CC associated with inappropriate SNK activity or levels. Disease states that  
 CC can be diagnosed include proliferative diseases such as leukemia, solid  
 CC tumour cancers and metastases, chronic inflammatory proliferative  
 CC diseases such as psoriasis and rheumatoid arthritis, proliferative  
 CC cardiovascular diseases such as restenosis, proliferative ocular  
 CC disorders such as diabetic retinopathy and benign hyperproliferative  
 CC diseases such as haemangiomas. The polynucleotides can be used as  
 CC hybridisation probes for cDNA and genomic DNA or as primers for a nucleic  
 CC acid amplification (PCR) reaction, to isolate full-length cDNAs and  
 CC genomic clones encoding polypeptides of this invention and to isolate  
 CC cDNA and genomic clones of other genes which have a high sequence  
 CC similarity to the SNK coding sequence. The differences between cDNA and  
 CC genomic sequences can be observed and therefore mutations detected. Any  
 CC mutations may then be attributed to likely causative agents of disease.  
 CC The nucleotide sequences are also useful for chromosome identification

XX Sequence 685 AA;

Query Match 11.0%; Score 560.5; DB 2; Length 685;





**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 13:58:47 ; Search time 1133 Seconds

(without alignments)  
11340.902 Million cell updates/sec

Title: US-10-026-021-1

Perfect score: 2913

Sequence: 1 atggcgactgcgcgcggga.....cgacctcaatttcattga 2913

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

```
Database :
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*
37: gb_ov:*
38: gb_ov:*
39: gb_ov:*
40: gb_ov:*
41: gb_ov:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2911.4	99.9	3225	9 BC036023	BC036023 Homo sapi
2	2911.4	99.9	3225	9 AR338789	AR338789 Sequence
3	2906.6	99.8	3092	9 HSTTPKSAK	Y13115 Homo sapien
4	2905	99.7	3331	9 AB006972	AB006972 Homo sapi
5	1879	64.5	3447	6 AR083569	AR083569 Sequence
6	1879	64.5	3447	6 I56871	I56871 Sequence 3
7	1879	64.5	3447	6 AX306014	AX306014 Sequence
8	1879	64.5	3447	10 M05SARA	L29479 Mus muscu
9	1874.2	64.3	3473	10 BC026785	BC026785 Mus muscu
10	1724.6	59.2	3377	10 BC057940	BC057940 Mus muscu
11	1064.2	36.5	3554	5 BC060363	BC060363 Xenopus 1
12	1025	35.2	57057	9 AC107053	AC107053 Homo sapi
13	1025	35.2	15342	2 AL591646	AL591646 Homo sapi
14	970.4	33.3	1600	6 AR083570	AR083570 Sequence
15	970.4	33.3	1600	6 I56872	I56872 Sequence 5
16	970.4	33.3	1600	6 M05SARB	L29480 Mus muscu
17	924.2	31.7	1453	6 AR083568	AR083568 Sequence
18	924.2	31.7	1453	6 I56870	I56870 Sequence 1
19	731.8	25.1	233181	2 AC111558	AC111558 Rattus no
20	683.4	23.5	158533	2 AC146980	AC146980 Mus muscu
21	619.6	21.3	201286	2 AC102228	AC102228 Mus muscu
22	619.6	21.3	232353	2 AC127590	AC127590 Mus muscu
23	373	12.8	2453	5 BC045434	BC045434 Danio rer
24	329.2	11.3	331	6 AX339849	AX339849 Sequence
25	313.8	10.8	972	10 BC051483	BC051483 Mus muscu
26	282.2	9.7	2574	3 AP106952	AP106952 Drosophi
27	280.6	9.6	3384	3 BT004846	BT004846 Drosophi
28	268.2	9.2	149008	9 AC007991	AC007991 Homo sapi
29	239	8.2	115884	2 AC017581	AC017581 Drosophi
30	239	8.2	168058	3 AC010049	AC010049 Drosophi
31	239	8.2	174920	3 AC010701	AC010701 Drosophi
32	239	8.2	314906	3 AE003594	AE003594 Drosophi
33	186.4	6.4	66081	2 AC083920	AC083920 Homo sapi
34	160.8	5.5	1573	3 AK113631	AK113631 Clona int
35	154.2	5.3	381	6 AR270086	AR270086 Sequence
36	146.8	5.0	3923	9 AK128691	AK128691 Homo sapi
37	137.8	4.7	2662	5 BC042254	BC042254 Xenopus 1
38	137.8	4.7	3589	5 AP357840	AP357840 Xenopus 1
39	136	4.7	2180	9 HS085755	U85755 Human betum
40	136	4.7	2375	9 AK098163	AK098163 Homo sapi
41	135.6	4.7	2972	9 AF059617	AF059617 Homo sapi
42	135	4.6	2381	5 X1U58205	U58205 Xenopus lae
43	135	4.6	2772	6 AX305275	AX305275 Sequence
44	135	4.6	2772	10 M05SNK	M96163 Mus muscu
45	135	4.6	2804	10 BC034513	BC034513 Mus muscu

## ALIGNMENTS

```
RESULT 1
BC036023 3225 bp mRNA PRI 12-NOV-2003
LOCUS Homo sapiens serine/threonine kinase 18, mRNA (CDNA clone MGC:33045
DEFINITION IMAGE:5273226), complete cds.
ACCESSION BC036023
VERSION BC036023.1 GI:23243308
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3225)
Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
Klaunig,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
```



Db 903 AATAAGATGATGACATTAAGAAATGCAATCTTTTGTCAATAGAGCCAAAGAC 962  
Qy 721 CTTATTACACAGTACTTGTGTAAGAAATCCAGCAGATGCTTTAACTGTCTTCAATATG 780  
Db 963 CTTATTACACAGTACTTGTGTAAGAAATCCAGCAGATGCTTTAACTGTCTTCAATATG 1022  
Qy 781 GACCAATCCTTTATGTCGGAAATTCCTTCAACAAAAGTAAGATTAGGAATGAGAA 840  
Db 1023 GACCAATCCTTTATGTCGGAAATTCCTTCAACAAAAGTAAGATTAGGAATGAGAA 1082  
Qy 841 GACTCAATGATAGTGGGATGCGACAAATTTCTACTGCAATTAACAGCTTCTTCACTAC 900  
Db 1083 GACTCAATGATAGTGGGATGCGACAAATTTCTACTGCAATTAACAGCTTCTTCACTAC 1142  
Qy 901 AGTATAGTGTAGTATTATTGACAAAAGACATTTGATGTCAGCCACTCCCAAT 960  
Db 1143 AGTATAGTGTAGTATTATTGACAAAAGACATTTGATGTCAGCCACTCCCAAT 1202  
Qy 961 AAAATGACGTATTTCCAAAGAAATAAGTCAACTGATTTTCTTCAAGAGATGGA 1020  
Db 1203 AAAATGACGTATTTCCAAAGAAATAAGTCAACTGATTTTCTTCAAGAGATGGA 1262  
Qy 1021 AACAGTTTTATATCTCAGTGGGAAATCAAGAAACAGTAAATAGTGAAGGGAGAGTA 1080  
Db 1263 AACAGTTTTATATCTCAGTGGGAAATCAAGAAACAGTAAATAGTGAAGGGAGAGTA 1322  
Qy 1081 ATTCAAGATGCAAGAAAGGCCACATTTCCATACCTTGTAGAGCTTAATCTCTGAT 1140  
Db 1323 ATTCAAGATGCAAGAAAGGCCACATTTCCATACCTTGTAGAGCTTAATCTCTGAT 1382  
Qy 1141 AGATCGGACCTTCTAATGTGCTCAAGCAAAAACATATCAATGGAAGATGTCAC 1200  
Db 1383 AGATCGGACCTTCTAATGTGCTCAAGCAAAAACATATCAATGGAAGATGTCAC 1442  
Qy 1201 TCAGCAAGAAATGCTTCACTGTCCAAAGATCAGAGAGTGAAGAAATGAGAGAGTAC 1260  
Db 1443 TCAGCAAGAAATGCTTCACTGTCCAAAGATCAGAGAGTGAAGAAATGAGAGAGTAC 1502  
Qy 1261 TCACCCACAGACAAACATGCTCAATTTTAACTTTTAAAGAAAGACATCTCAATAGT 1320  
Db 1503 TCACCCACAGACAAACATGCTCAATTTTAACTTTTAAAGAAAGACATCTCAATAGT 1562  
Qy 1321 TCTGATCTTTTGAAGACCTGATTAACATCAAGCACTCTCAATATCTTTTCCAGGA 1380  
Db 1563 TCTGATCTTTTGAAGACCTGATTAACATCAAGCACTCTCAATATCTTTTCCAGGA 1622  
Qy 1381 AAAATCCTTTTCAATTTGACAGCCGACACCTCAGACTGAAACCTTACAAACAGTGT 1440  
Db 1623 AAAATCCTTTTCAATTTGACAGCCGACACCTCAGACTGAAACCTTACAAACAGTGT 1682  
Qy 1441 GGGAAATCTGCAAAATTAATGCTCTATTAAGAAAACTACTGAATATGACAGCATCAGCCA 1500  
Db 1683 GGGAAATCTGCAAAATTAATGCTCTATTAAGAAAACTACTGAATATGACAGCATCAGCCA 1742  
Qy 1501 AACCCGGAATTCAGAGGCAATCCAGTTTGGAGAAAGACATCAAAAATGCTGGAAT 1560  
Db 1743 AACCCGGAATTCAGAGGCAATCCAGTTTGGAGAAAGACATCAAAAATGCTGGAAT 1802  
Qy 1561 GATACAAAAGTCAAAAAGAACTCTGATGCTTCTGATATGACATTCGTGTAACAGCAA 1620  
Db 1803 GATACAAAAGTCAAAAAGAACTCTGATGCTTCTGATATGACATTCGTGTAACAGCAA 1862  
Qy 1621 AATACATGAATAATATATGACTGCACTTCAACAGTAACCTGAGATATCCAAACAGATGT 1680  
Db 1863 AATACATGAATAATATATGACTGCACTTCAACAGTAACCTGAGATATCCAAACAGATGT 1922  
Qy 1681 GTTTTGGCTCAGATCTCTTCTGAAACAGAGCAACATAGGGGTATGAGCCACATGG 1740  
Db 1923 GTTTTGGCTCAGATCTCTTCTGAAACAGAGCAACATAGGGGTATGAGCCACATGG 1982  
Qy 1741 GGTATCAGAAATGATCAATTAAGAAAGATTAACATCTCCGTGGTCTCAACAGTTTAA 1800

Db 1983 GGTATCAGAAATGATCAATTAAGAAAGATTAACATCTCCGTGGTCTCAACAGTTTAA 2042  
Qy 1801 CCAATTCAGACAAAAACAAAAGCTGTGGAGAGATCTTGATTCAGAGAGGTGT 1860  
Db 2043 CCAATTCAGACAAAAACAAAAGCTGTGGAGAGATCTTGATTCAGAGAGGTGT 2102  
Qy 1861 GTGGAGCTTTGAAGAGATGATCATCTCAAGAAATATGTAAAGAAAGTTCTTCAGATATCT 1920  
Db 2103 GTGGAGCTTTGAAGAGATGATCATCTCAAGAAATATGTAAAGAAAGTTCTTCAGATATCT 2162  
Qy 1921 AGTATGAAATACGATCACTATTATTAATCAAAATGTGTGAGAGTTTCTCTGCT 1980  
Db 2163 AGTATGAAATACGATCACTATTATTAATCAAAATGTGTGAGAGTTTCTCTGCT 2222  
Qy 1981 GATAGCAACCTCAGCTCTGACAAATCAGATAGGTAACGCTTGAACAATTAACAGAA 2040  
Db 2223 GATAGCAACCTCAGCTCTGACAAATCAGATAGGTAACGCTTGAACAATTAACAGAA 2282  
Qy 2041 AAAATCTGGCAAAAATATCAATATGCTTCCAGTTTGTACAGCTTGTAAATCTAATCT 2100  
Db 2283 AAAATCTGGCAAAAATATCAATATGCTTCCAGTTTGTACAGCTTGTAAATCTAATCT 2342  
Qy 2101 CCAAAAATCACTTATTTTAAAGATATGCTAAATGCAATTTTGAAGAAATCTCTGCT 2160  
Db 2343 CCAAAAATCACTTATTTTAAAGATATGCTAAATGCAATTTTGAAGAAATCTCTGCT 2402  
Qy 2161 GCTGATTTGAGTTTGTGTTTATGATGGGTTAAATACAAAACAGAAATTTCAAT 2220  
Db 2403 GCTGATTTGAGTTTGTGTTTATGATGGGTTAAATACAAAACAGAAATTTCAAT 2462  
Qy 2221 CAGTGATTTGAAGAGACAGGAAAGTCTTACCTTTAAAGTGAAGTAAATATAC 2280  
Db 2463 CAGTGATTTGAAGAGACAGGAAAGTCTTACCTTTAAAGTGAAGTAAATATAC 2522  
Qy 2281 TTGAAGAGAGATTAATAATGATATGACATGCTAATGAGGTCATGTAATTTGTTA 2340  
Db 2523 TTGAAGAGAGATTAATAATGATATGACATGCTAATGAGGTCATGTAATTTGTTA 2582  
Qy 2341 GCACTGGAATCCATTAATTTCAAGAGAAAGAAAGAAAGTAAAGTCCCTTTTCCCA 2400  
Db 2583 GCACTGGAATCCATTAATTTCAAGAGAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2642  
Qy 2401 ATATCATGAGAAAGAAACCTGTAGTACTAGTTTCACTTAAGGCTTATCACTCTCT 2460  
Db 2643 ATATCATGAGAAAGAAACCTGTAGTACTAGTTTCACTTAAGGCTTATCACTCTCTCT 2702  
Qy 2461 TCTGTGATTCAAATTAATCCCAAGAGATGAGCATCTTCAACAGAAATGTCATGAT 2520  
Db 2703 TCTGTGATTCAAATTAATCCCAAGAGATGAGCATCTTCAACAGAAATGTCATGAT 2762  
Qy 2521 AGTGTGCTTCTTCCAAACAGGCAACATCTTAAATCCCTTATGTTTCAAAATGAAGA 2580  
Db 2763 AGTGTGCTTCTTCCAAACAGGCAACATCTTAAATCCCTTATGTTTCAAAATGAAGA 2822  
Qy 2581 CTTGGCTTCAACACTACAGCTTCTGGAACAGACATCTCTTAAATGCTTAAAGATGT 2640  
Db 2823 CTTGGCTTCAACACTACAGCTTCTGGAACAGACATCTCTTAAATGCTTAAAGATGT 2882  
Qy 2641 CTTCTTAAATCAGCAACATTTTGAATCTGTTTTGTGAATAATGTGTTGGGTACA 2700  
Db 2883 CTTCTTAAATCAGCAACATTTTGAATCTGTTTTGTGAATAATGTGTTGGGTACA 2942  
Qy 2701 CAGTTAATCTAGAGAGTGTGGGTTCAGTTTAAATGATGAGTCCCAAGTTGTGTCAG 2760  
Db 2943 CAGTTAATCTAGAGAGTGTGGGTTCAGTTTAAATGATGAGTCCCAAGTTGTGTCAG 3002  
Qy 2761 GCAGAGAGTCTTCTATCAGTTATACCTCAACAAATGTCAAACAACTAGTATGAGAA 2820  
Db 3003 GCAGAGAGTCTTCTATCAGTTATACCTCAACAAATGTCAAACAACTAGTATGAGAA 3062  
Qy 2821 AATGAAAAATTAACAGACTACATCAACAGAAATTAACAGTGTCTGTCTTCAATCTTTTG 2880  
Db 3063 AATGAAAAATTAACAGACTACATCAACAGAAATTAACAGTGTCTGTCTTCAATCTTTTG 3122

Qy	2881	ATGTTTCTAATCCGACTCCTAATTTTCATGGA	2913	
Db	3123	ATGTTTCTAATCCGACTCCTAATTTTCATGGA	3155	
RESULT 2				
LOCUS	AR338789	3937 bp	DNA	linear
DEFINITION	Sequence 280 from patent US 656962.			
ACCESSION	AR338789			
VERSION	AR338789.1	GI:33725646		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 3937)			
TITLE	Tang, Y.T., Zhou, P. and Dmanac, R.T.			
JOURNAL	Nucleic acids and polypeptides			
FEATURES	Patent: US 656962-A 280 27-MAY-2003;			
source	Location/Qualifiers			
	1..3937			
ORIGIN	/organism="unknown"			
	/mol_type="genomic DNA"			
Query Match	99.9%;	Score 2911.4;	DB 6;	Length 3937;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches	0;
Matches 2912;	Conservative	0;	Indels	0;
			Gaps	0;
Qy	1	ATGCGACCTCGATCGGGGAGAAATCCAGGATTTTAAAGTTGAAATCTGCTGTGTA	60	
Db	334	ATGCGACCTCGATCGGGGAGAAATCCAGGATTTTAAAGTTGAAATCTGCTGTGTA	393	
Qy	61	GGATCATTTTGTGTGTCTTACAGAGCTGATCCATTCACTGGTTTGAAGTTGCAATC	120	
Db	394	GGATCATTTTGTGTGTCTTACAGAGCTGATCCATTCACTGGTTTGAAGTTGCAATC	453	
Qy	121	AAATGATGATGAAGAAAGCCATGTACAAAGCAGAAATGTACAGAGTCCAAATGAG	180	
Db	454	AAATGATGATGAAGAAAGCCATGTACAAAGCAGAAATGTACAGAGTCCAAATGAG	513	
Qy	181	GTGAATATCATTTGCCAATTGAAACATCTTCTATCTTGGAGCTTTTAACTATTTGAA	240	
Db	514	GTGAATATCATTTGCCAATTGAAACATCTTCTATCTTGGAGCTTTTAACTATTTGAA	573	
Qy	241	GATGACATTTATGTATCTGTATCTGTATGAAATGTGCCATATATGAGAAATGAA	300	
Db	574	GATGACATTTATGTATCTGTATCTGTATGAAATGTGCCATATATGAGAAATGAA	633	
Qy	301	CTAAAGAAATGAGTGAACCCCTTCTCAGAAATGAACTCGACATTCATGACACAGATC	360	
Db	634	CTAAAGAAATGAGTGAACCCCTTCTCAGAAATGAACTCGACATTCATGACACAGATC	693	
Qy	361	ATCAGAGGATGTGTATCTTCAATTCATGATGTATCTAATCAGGAGCTTCACTTTCT	420	
Db	694	ATCAGAGGATGTGTATCTTCAATTCATGATGTATCTAATCAGGAGCTTCACTTTCT	753	
Qy	421	AACCTCTACTGACTCGTAAATATGAAATCAAGATTTGCTGATTTTGGGCTGGCACTCAA	480	
Db	754	AACCTCTACTGACTCGTAAATATGAAATCAAGATTTGCTGATTTTGGGCTGGCACTCAA	813	
Qy	481	CTGAATATGCGACATGAAAGCCTATACATTTATGTGAACTCTTAACTTACATTTCA	540	
Db	814	CTGAATATGCGACATGAAAGCCTATACATTTATGTGAACTCTTAACTTACATTTCA	873	
Qy	541	GAAATTCGCACTCGAAGTGCATGAGCTTGAATCTGATGTTTGTCCCTGGGCTGTATG	600	
Db	874	GAAATTCGCACTCGAAGTGCATGAGCTTGAATCTGATGTTTGTCCCTGGGCTGTATG	933	
Qy	601	TTTATATCATTTATCTTATCGGGAACCACTTGTGACATCTGACACAGTCAAGAACATTA	660	
Db	934	TTTATATCATTTATCTTATCGGGAACCACTTGTGACATCTGACACAGTCAAGAACATTA	993	
Qy	661	AATTAAGTATATTGGCAGATTAATGAATGCCATCTTTTGTCAATAGAGCCAAAGAC	720	
Db	994	AATTAAGTATATTGGCAGATTAATGAATGCCATCTTTTGTCAATAGAGCCAAAGAC	1013	
Qy	721	CTTATTCACAGTTACTCTGTGAAATTCAGAGATCTGTTAAGTCTGTCTTCAATATG	780	
Db	1054	CTTATTCACAGTTACTCTGTGAAATTCAGAGATCTGTTAAGTCTGTCTTCAATATG	1113	
Qy	781	GACCATCTTTTATGTCCGAAATTTCTTCAACAAAGATTAAGATTAGAACTGTGAA	840	
Db	1114	GACCATCTTTTATGTCCGAAATTTCTTCAACAAAGATTAAGATTAGAACTGTGAA	1113	
Qy	841	GACTCAATTGATAGTGGGATGCGACCAATTTCTATCGAATTAACAGCTTCTTCAATG	900	
Db	1174	GACTCAATTGATAGTGGGATGCGACCAATTTCTATCGAATTAACAGCTTCTTCAATG	1223	
Qy	901	AGTATTAAGGTGATGTTTATTTTGAACAAAGAGACTTTTGAATGTGTGACCACTCCAAAT	960	
Db	1234	AGTATTAAGGTGATGTTTATTTTGAACAAAGAGACTTTTGAATGTGTGACCACTCCAAAT	1223	
Qy	961	AAATGACTGTATTTTCCAAAGATTAAGATTCACTGATTTTCTTCTTCAAGAGATGGA	1020	
Db	1294	AAATGACTGTATTTTCCAAAGATTAAGATTCACTGATTTTCTTCTTCAAGAGATGGA	1359	
Qy	1021	AACAGTTTATATCTCACTGCGGAAATCAAGAAACAGTAATATGGAAGGGAAGTA	1080	
Db	1354	AACAGTTTATATCTCACTGCGGAAATCAAGAAACAGTAATATGGAAGGGAAGTA	1411	
Qy	1081	ATTCAAGATGCAAGAAAGGCCACATCTGTGATACCTGTGTAGCTTATTCCTCTGAT	1140	
Db	1414	ATTCAAGATGCAAGAAAGGCCACATCTGTGATACCTGTGTAGCTTATTCCTCTGAT	1477	
Qy	1141	AGATGTGCACTTCTAATATGTCAGTCTTCAAGCAAAAATATTAATATGAAAGATGTCAC	1200	
Db	1474	AGATGTGCACTTCTAATATGTCAGTCTTCAAGCAAAAATATTAATATGAAAGATGTCAC	1533	
Qy	1201	TCAGCAAGAAATGCTTTCAGTGTCCAAAGATCAGAGAGAGTGAAATGAAGAGAGTAC	1260	
Db	1534	TCAGCAAGAAATGCTTTCAGTGTCCAAAGATCAGAGAGAGTGAAATGAAGAGAGTAC	1599	
Qy	1261	TCACCCACAGCAACATGCCACATTTTAACTTTTAAAGAAAGACATCCAGTAGT	1320	
Db	1594	TCACCCACAGCAACATGCCACATTTTAACTTTTAAAGAAAGACATCCAGTAGT	1655	
Qy	1321	TCTGATCTTTTGAAGAACTGTATTAACATCAAGCACTCTCCAAATCATCTTGTCCAGGA	1380	
Db	1654	TCTGATCTTTTGAAGAACTGTATTAACATCAAGCACTCTCCAAATCATCTTGTCCAGGA	1711	
Qy	1381	AAAATCTCTTTTCAATTTGCAAGCCGACACTTCAAGCTGAAACCGTACAAAGTGT	1441	
Db	1714	AAAATCTCTTTTCAATTTGCAAGCCGACACTTCAAGCTGAAACCGTACAAAGTGT	1777	

1741 GGTATACAGATCGTACATTAAAGACATTAATCTCCGTTGGTCTCAGAGTTAA 1800  
 2074 GATTATCAATCTGATTAAGAGACATTAATCTCCGTTGGTCTCAGAGTTAA 2133  
 1801 CCAATCAGACAGAAAACCAAAAAGGTGTGTGAGCATCTTGATTCAGAGAGGTGT 1860  
 2134 CCAATCAGACAGAAAACCAAAAAGGTGTGTGAGCATCTTGATTCAGAGAGGTGT 2193  
 1861 GTGAGCTGTGTAAGAGATGATCTCAAGAAATATGTGAAAAGAGTTCTTCAAGATCT 1920  
 2194 GTGAGCTGTGTAAGAGATGATCTCAAGAAATATGTGAAAAGAGTTCTTCAAGATCT 2253  
 1921 AGTGAATGAATATGATCTCAATTTATTTATTCCAAAATGTGTGAGAGTTTCTCTTGTCT 1980  
 2254 AGTGAATGAATATGATCTCAATTTATTTATTCCAAAATGTGTGAGAGTTTCTCTTGTCT 2313  
 1981 GATAAGCCACCTCACTCTGACCAATCAGTAGGTACAGCTTTGACAAATTTACAGAA 2040  
 2314 GATAAGCCACCTCACTCTGACCAATCAGTAGGTACAGCTTTGACAAATTTACAGAA 2373  
 2041 AATATCTGGCGAAATATCAATATGCTTCAGAGTTGTATCAGCTTTGAATCTAATCT 2100  
 2374 AATATCTGGCGAAATATCAATATGCTTCAGAGTTGTATCAGCTTTGAATCTAATCT 2433  
 2101 CCAAAATCACTTATTTTCAAGATATGCTAAATGCAATTTTGTGAGAAATCTCTCTGT 2160  
 2434 CCAAAATCACTTATTTTCAAGATATGCTAAATGCAATTTTGTGAGAAATCTCTCTGT 2493  
 2161 GCTGATTTTGAAGTTTGTATTTATGATGAGGTTAAATACAAAACAGAAATTTTCAAT 2220  
 2494 GCTGATTTTGAAGTTTGTATTTATGATGAGGTTAAATACAAAACAGAAATTTTCAAT 2553  
 2221 CAGGTGATTTGAAAAGACAGAGGAAGTTTACACTTTAAAGTGAAGTGAATATAC 2280  
 2554 CAGGTGATTTGAAAAGACAGAGGAAGTTTACACTTTAAAGTGAAGTGAATATAC 2613  
 2281 TTGAAAGAGAGATTAATAATGTATATGACCAATGCTAAATGAGGTCATGATTTGTTA 2340  
 2614 TTGAAAGAGAGATTAATAATGTATATGACCAATGCTAAATGAGGTCATGATTTGTTA 2673  
 2341 GCACTGGAATCCATATATTTCAAGAGAGAAAGAACTAGAGAGTCTCCCTTTTCCCA 2400  
 2674 GCACTGGAATCCATATATTTCAAGAGAGAAAGAACTAGAGAGTCTCCCTTTTCCCA 2733  
 2401 AATATCATTGAAAGAAACCTGTGATGATCTTCACTTAAAGCTTTATCTCTCTCT 2460  
 2734 AATATCATTGAAAGAAACCTGTGATGATCTTCACTTAAAGCTTTATCTCTCTCTCT 2793  
 2461 TCTGTGAGATTTCAAAATTAACCAACGAGATGAGATCTTTCAAGAGATGATGAT 2520  
 2794 TCTGTGAGATTTCAAAATTAACCAACGAGATGAGATCTTTCAAGAGATGATGAT 2853  
 2521 AGTGTGCTTTCTTCAACAGAGGACCAATCTTAAATCTCTTAAATGTTAAGTGAAG 2580  
 2854 AGTGTGCTTTCTTCAACAGAGGACCAATCTTAAATCTCTTAAATGTTAAGTGAAG 2913  
 2581 CTTGGTCTTAACTACAGCTTCTGGAACAGACATCTTTTAAATGTTTAAAGATGTT 2640  
 2914 CTTGGTCTTAACTACAGCTTCTGGAACAGACATCTTTTAAATGTTTAAAGATGTT 2973  
 2641 CTTTCTTAATCAAGCAAACTTTTGAATCTGTTTTGTGAAAATGTGTGTTGGGCTACA 2700  
 2974 CTTTCTTAATCAAGCAAACTTTTGAATCTGTTTTGTGAAAATGTGTGTTGGGCTACA 3033  
 2701 CAGTTAACTAGTGAAGCTGTGTGTTCAATTTAATGATGAGTCCAGTTGTGTGAG 2760  
 3034 CAGTTAACTAGTGAAGCTGTGTGTTCAATTTAATGATGAGTCCAGTTGTGTGAG 3093  
 2761 GAGAGAGTGTCTTCAATCAATTTAATCTCAAAATGTGTGAACATAGATTTGAGAA 2820  
 3094 GAGAGAGTGTCTTCAATCAATTTAATCTCAAAATGTGTGAACATAGATTTGAGAA 3153  
 2821 AATGAAATTTACAGATCAATCAAAAGAAATTTACAGATGTGTCTTCAATCTTTTG 2880

Db 3154 AATGAAATTTACAGAGATCAATCAAGAAATTAAGATGTGTCTTCAATCTTTTG 3213  
 Qy 2881 AATGTTTCTTAATCCGACTCTTAATTTTCAATGA 2913  
 Db 3214 AATGTTTCTTAATCCGACTCTTAATTTTCAATGA 3246

RESULT 3  
 HSTPCKSAK  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

HSTPCKSAK 3092 bp mRNA linear PRI 27-MAY-1997  
 Homo sapiens mRNA for serine/threonine protein kinase SAK.  
 Y13115  
 Y13115.1 GI:2125813  
 serine/threonine protein kinase.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
 1  
 Karn, T., Holtrich, U., Wolf, G., Hock, B., Strubhardt, K. and  
 Knebelmann-Waigmann, H.  
 Human SAK related to the PKC/polo family of cell cycle kinases  
 shows high mRNA expression in testis  
 Oncol. Rep. 4, 505-510 (1997)  
 2 (bases 1 to 3092)  
 Karn, T.  
 Direct Submission  
 Submitted (09-MAY-1997) T. Karn, Chemotherapeutisches  
 Forschungsinstitut, Georg-Speyer-Haus, Paul-Ehrlich-Strasse 42-44,  
 D-60596 Frankfurt, FRG

FEATURES  
 source  
 1..3092  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="X3"  
 /tissue\_type="lung"  
 1..3092  
 /gene="SAK"  
 141..3053  
 /gene="SAK"  
 /codon\_start=1  
 /product="serine/threonine protein kinase"  
 /protein\_id="CAA73575.1"  
 /db\_xref="GI:2125814"  
 /db\_xref="GOA:O00444"  
 /db\_xref="SPTREMBL:O00444"  
 /translation="MATCIGEKIEDPKFKNGLGKSPAGVYRAISIHGIEVALIKMID  
 KKAMKAGVORVKNKVIKHOLKAPSLILEVNFEDSNVYLVLEMCHGEMRYLK  
 NKVPSERENARHPNHOITTMGLYHSHGILHRDITSLTLTRMMNITKIDRGALAT  
 LMPHEKRYTLCTGPNVISPETATRSAGLESVDVSLGCMFTLLIGRPEDTIVAN  
 TLNKVLVDYEMPTLSIEADLIHQLRNPADRLSLSVLDHPMSNSSTKSDL  
 GTVEDSIDSGATISTAITASSTSIISGLFPRKRLIGQPLPKMKTGPKKSSDF  
 SSSGDSFYIOWGNETSNGRGVIODAEPRPSRYLRAYSIDRSSTSSQOAK  
 TTYMRCHSAEMLSVSKSGSENEREYSPTDMNATINPFEKSTSSSGSEPRPNN  
 QLSNHLCEGKTPPPADPTPOTRTVQAFGLQINAHKRTTEIDISIPANDPQCHP  
 DLQKTSIKNAWTDIVKKNDSADNASHVKQNTKMTALSKBEIIIOECVPSDP  
 LSEOKRTGEMPPMYONKTLASISPLVAHLKPIROTKKAVVSIIDSEVCVELV  
 KEVASOEYKVELVQISDQNTTIIYVNGRGPPADRPSPPTDMSISYSPNLBEKY  
 WKVOYASRFVLVRSKSPKTIYPTRYAKCLIMENSPGDPFVWFYDGKLIKHTEDPI  
 QVIERKGSYTIKSSSVNLSKEIKMPEDANBGRHICLALESIISEERETRAAF  
 PFIIGRKPGSTSSPKALSPPSVDNSTPTDRASPRNRMHSAASPTQAPILNBSMV  
 TNEGGLTTLTASGTDISSNLSKDCPKSAQLKSVFVKNVGNATULTSGAVVQFSDG  
 SOLVVOAGVSSISYSPNGQTRRYENEKLPDIYIKQKQCSSIILMFSNPENPH"

ORIGIN  
 Query Match 99.8%; Score 2906.6; DB 9; Length 3092;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2909; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AATGCACTGATCGGAGAGAAATCGAGATTTTAAAGTTGAATCTGCTGTGA 60

Dh 141 ATGGCGACTGTCATCGGGGAGAAATCGAGATTTTAAAGTTTGAATCTGCTGGTAA 200  
Qy 61 GGAATCTTGGCTGGTGTCTACAGAGCTGAGTCCATTCACTGGTGGAAATGGTCAATC 120  
Dh 201 GGAATCTTGGCTGGTGTCTACAGAGCTGAGTCCATTCACTGGTGGAAATGGTCAATC 260  
Qy 121 AAAATGATAGTAAAGAACCATGTACAAAGCAGGAAATGTACAGAGTCCAAATAG 180  
Dh 261 AAAATGATAGTAAAGAACCATGTACAAAGCAGGAAATGTACAGAGTCCAAATAG 320  
Qy 181 GTGAAATCATTTGCCAATTTGAACATCTTCTATCTTGGAGCTTTATTAATTTGAA 240  
Dh 321 GTGAAATCATTTGCCAATTTGAACATCTTCTATCTTGGAGCTTTATTAATTTGAA 380  
Qy 241 GATACCAATTAATGTGTATCTGGTATTAGAAATGTGCCATAATGAGAAATGAAGGAT 300  
Dh 381 GATACCAATTAATGTGTATCTGGTATTAGAAATGTGCCATAATGAGAAATGAAGGAT 440  
Qy 301 CTAAAGATTAAGTGAACCCCTTCTCAGAAATGAAAGCTGACACTTCATGACCAAGATC 360  
Dh 441 CTAAAGATTAAGTGAACCCCTTCTCAGAAATGAAAGCTGACACTTCATGACCAAGATC 500  
Qy 361 ATCAAGAGATGTGTATCTTCAATCTCATGTATTAATCAACCGGAGCTTCACTTTCT 420  
Dh 501 ATCAAGAGATGTGTATCTTCAATCTCATGTATTAATCAACCGGAGCTTCACTTTCT 560  
Qy 421 AACCTCCATGACTCGTAAATATGAACATCAAGATGCTGATTTGGGCTGGCACTCA 480  
Dh 561 AACCTCCATGACTCGTAAATATGAACATCAAGATGCTGATTTGGGCTGGCACTCA 620  
Qy 481 CTGAAATGCGACATGAAGAACATTAATATGAGAACTCTTAATCAATTTACCA 540  
Dh 621 CTGAAATGCGACATGAAGAACATTAATATGAGAACTCTTAATCAATTTACCA 680  
Qy 541 GAAATGCGACATGAAGAACATGAGCTTGAATCTGATGTTGGTCCCTGGCTGATG 600  
Dh 681 GAAATGCGACATGAAGAACATGAGCTTGAATCTGATGTTGGTCCCTGGCTGATG 740  
Qy 601 TTTTATCATTTACTTATCGGAGAACACCTTGGACATGACAGTCAAGAACATTA 660  
Dh 741 TTTTATCATTTACTTATCGGAGAACACCTTGGACATGACAGTCAAGAACATTA 800  
Qy 661 AATTAAGTATGATGAGCAATTAAGAAATGCCATCTTTTGTCAATAGAGCCAAAGAC 720  
Dh 801 AATTAAGTATGATGAGCAATTAAGAAATGCCATCTTTTGTCAATAGAGCCAAAGAC 860  
Qy 721 CTATTCACCAAGTACTTGTGTAGAAATCAAGCAAGATGCTTAAATCTCTTCAATATG 780  
Dh 861 CTATTCACCAAGTACTTGTGTAGAAATCAAGCAAGATGCTTAAATCTCTTCAATATG 920  
Qy 781 GACCATCTTTTATGTCCGAAATCTTCAACAAAGTAAAGATTTGGAACGTGTGAA 840  
Dh 921 GACCATCTTTTATGTCCGAAATCTTCAACAAAGTAAAGATTTGGAACGTGTGAA 980  
Qy 841 GACTCAATGATAGGAGGAGATGCCCAATTTTACTGCAATTAAGCTTCTTCCAGTACC 900  
Dh 981 GACTCAATGATAGGAGGAGATGCCCAATTTTACTGCAATTAAGCTTCTTCCAGTACC 1040  
Qy 901 AGTATAGTGTAGTATTATTTGACAAAGAGAGACTTTTGTATGAGGAGCCCTCCCAAT 960  
Dh 1041 AGTATAGTGTAGTATTATTTGACAAAGAGAGACTTTTGTATGAGGAGCCCTCCCAAT 1100  
Qy 961 AAAATGACTGATTTTCCAAAGATTAAGTTCATCTGATTTTCTTCCAGAGATGGA 1020  
Dh 1101 AAAATGACTGATTTTCCAAAGATTAAGTTCATCTGATTTTCTTCCAGAGATGGA 1160  
Qy 1021 AACAGTTTATTACTCAGTGGGAGAAATCAAGAAACAGTAAATAGTGAAGGGAGTGA 1080  
Dh 1161 AACAGTTTATTACTCAGTGGGAGAAATCAAGAAACAGTAAATAGTGAAGGGAGTGA 1220  
Qy 1081 ATTCAAGATGAGAAAGGCGCAATTTCTGATACCTTGTAGAGCTTATTCCTGAT 1140  
Dh 1221 ATTCAAGATGAGAAAGGCGCAATTTCTGATACCTTGTAGAGCTTATTCCTGAT 1280

Qy 1141 AGATCTGGGACTTCTAATAGTCAAGTCTCAAGCAAAACATATACATGAAAGATGTAC 1200  
Dh 1281 AGATCTGGGACTTCTAATAGTCAAGTCTCAAGCAAAACATATACATGAAAGATGTAC 1340  
Qy 1201 TCAGCAAAATGCTTTCAGTGTCCAAGATGAGAGAGGAGTAAATGAAAGAGGTAC 1260  
Dh 1341 TCAGCAAAATGCTTTCAGTGTCCAAGATGAGAGAGGAGTAAATGAAAGAGGTAC 1400  
Qy 1261 TCACCCACAGACAAATGCCCAACTTTTAACTTCTTAAAGAAAGACATCCAGTAT 1320  
Dh 1401 TCACCCACAGACAAATGCCCAACTTTTAACTTCTTAAAGAAAGACATCCAGTAT 1460  
Qy 1321 TCTGATCTTTTGAAGAAGCTGATTAACATCAAGCACTCCATCATCTTTGTCCAGA 1380  
Dh 1461 TCTGATCTTTTGAAGAAGCTGATTAACATCAAGCACTCCATCATCTTTGTCCAGA 1520  
Qy 1381 AAAATCTCTTTTCCATTTTGCAGACCCGACCTCAGACTGAACCTGATCAACAGTGT 1440  
Dh 1521 AAAATCTCTTTTCCATTTTGCAGACCCGACCTCAGACTGAACCTGATCAACAGTGT 1580  
Qy 1441 GGAATCTGCAAAATTAATGCTCAATTAAGAAACCTAAGTATGACAGATGACCCA 1500  
Dh 1581 GGAATCTGCAAAATTAATGCTCAATTAAGAAACCTAAGTATGACAGATGACCCA 1640  
Qy 1501 AACCGGACTTCCAGGGCCATCCAGATTTGAGAAAGACATCAAAATATGCTGGACT 1560  
Dh 1641 AACCGGACTTCCAGGGCCATCCAGATTTGAGAAAGACATCAAAATATGCTGGACT 1700  
Qy 1561 GATACAAATGCAAAAGAACTCGATGCTTGTGAATGCAATTTGTAAACAGCA 1620  
Dh 1701 GATACAAATGCAAAAGAACTCGATGCTTGTGAATGCAATTTGTAAACAGCA 1760  
Qy 1621 AATACATGAATATATGACTGCACTTCAAGTAACTGAGATTAATCCAAAGAAATG 1680  
Dh 1761 AATACATGAATATATGACTGCACTTCAAGTAACTGAGATTAATCCAAAGAAATG 1820  
Qy 1681 GTTTTGGCTCAAGTCCCTCTTCTGAAACAGACAACTAAGGGGTATGAGCCACCATG 1740  
Dh 1821 GTTTTGGCTCAAGTCCCTCTTCTGAAACAGACAACTAAGGGGTATGAGCCACCATG 1880  
Qy 1741 GGTATACGAATGTATCAATTAAGAACATTAATCTCCGTGGTGTCTCAAGGTTAA 1800  
Dh 1881 GGTATACGAATGTATCAATTAAGAACATTAATCTCCGTGGTGTCTCAAGGTTAA 1940  
Qy 1801 CCAATCAGACAGAAACCAAAAGGCTGTGTGAGCATCTTGATTCAGAGAGGTGTG 1860  
Dh 1941 CCAATCAGACAGAAACCAAAAGGCTGTGTGAGCATCTTGATTCAGAGAGGTGTG 2000  
Qy 1861 GTGAGCTTGTAAAGAGATGATCTCAAGAAATGTGAAGAAATGCTTCAGATATCT 1920  
Dh 2001 GTGAGCTTGTAAAGAGATGATCTCAAGAAATGTGAAGAAATGCTTCAGATATCT 2060  
Qy 1921 AGTATGAGAAATACATCTATTTATTAATCCAAATGCTGAGAGGTTTCTCTTGTCT 1980  
Dh 2061 AGTATGAGAAATACATCTATTTATTAATCCAAATGCTGAGAGGTTTCTCTTGTCT 2120  
Qy 1981 GATAGACACCTCTCACTAGCAACATCAGTAGTACAGCTTTGAAATTAACAGAA 2040  
Dh 2121 GATAGACACCTCTCACTAGCAACATCAGTAGTACAGCTTTGAAATTAACAGAA 2180  
Qy 2041 AAAATCTGGCGAAATATCATATGCTTCCAGGTTTGTACAGCTTGAAGATCTAAATCT 2100  
Dh 2181 AAAATCTGGCGAAATATCATATGCTTCCAGGTTTGTACAGCTTGAAGATCTAAATCT 2240  
Qy 2101 CCCAAATCACTTATTTTAAAGATATGCTTAAAGCAATTTGATGAGAAATTCCTGCT 2160  
Dh 2241 CCCAAATCACTTATTTTAAAGATATGCTTAAAGCAATTTGATGAGAAATTCCTGCT 2300  
Qy 2161 GCTGATTTTGAAGGTTTGTATTAATGATGGGTAAATTAACAACAAGAGATTTCA 2220  
Dh 2301 GCTGATTTTGAAGGTTTGTATTAATGATGGGTAAATTAACAACAAGAGATTTCA 2360





```
QY 541 GAAATGGCACTCGAAGTGCAATGGCCTTGAATCTGAATGTTTGGTCCCTGGGCTGTATG 600
DB 729 GAAATGGCACTCGAAGTGCAATGGCCTTGAATCTGAATGTTTGGTCCCTGGGCTGTATG 788
QY 601 TTTTATACATTACTTATCGGAGAACCACTTCGACATGACACAGTCAAGAACATTA 660
DB 789 TTTTATACATTACTTATCGGAGAACCACTTCGACATGACACAGTCAAGAACATTA 848
QY 661 AATAAAGATATGAGGAGATTAAGAAATGCACTTTTGTCAATATAGGCGCAAGAC 720
DB 849 AATAAAGATATGAGGAGATTAAGAAATGCACTTTTGTCAATATAGGCGCAAGAC 908
QY 721 CTATTCACGAGTACTTCGTAGAAATCCAGCAGATGTTTAACTGTCTTCAGTATG 780
DB 909 CTATTCACGAGTACTTCGTAGAAATCCAGCAGATGTTTAACTGTCTTCAGTATG 968
QY 781 GACCATCTTTTATGTCCCGAAATTTCTCAACAAAAGTAAAGATTTAGGAACGTGGAA 840
DB 969 GACCATCTTTTATGTCCCGAAATTTCTCAACAAAAGTAAAGATTTAGGAACGTGGAA 1028
QY 841 GACTCAATGATAGTGGGCAATGCCAATTTCTACCTGCAATTAACGTTTCTTCCAGTACC 900
DB 1029 GACTCAATGATAGTGGGCAATGCCAATTTCTACCTGCAATTAACGTTTCTTCCAGTACC 1088
QY 901 AGTATAGTGTAGTATTTTGAACAAAAGAACCTTTGATGTGAGCCACTCCCAAT 960
DB 1089 AGTATAGTGTAGTATTTTGAACAAAAGAACCTTTGATGTGAGCCACTCCCAAT 1148
QY 961 AAATAGCTGTATTTCCAAAGAAATTAAGTTCAACGATTTTCTTCTTCAGAGATGGA 1020
DB 1149 AAATAGCTGTATTTCCAAAGAAATTAAGTTCAACGATTTTCTTCTTCAGAGATGGA 1208
QY 1021 AACAAGTTTTTAACTCAGTGGGGAATCAAGAAACGATATATGAGGAGGGAAGATA 1080
DB 1209 AACAAGTTTTTAACTCAGTGGGGAATCAAGAAACGATATATGAGGAGGGAAGATA 1268
QY 1081 ATTCAAGATGCAAGAAAGGCCACATTTCTGATACCTTGTAGAGCTTATTTCTCTGAT 1140
DB 1269 ATTCAAGATGCAAGAAAGGCCACATTTCTGATACCTTGTAGAGCTTATTTCTCTGAT 1328
QY 1141 AGATCTGGCACTTCTATATAGTCACTCAAGCAAAAACATATTAACATGGAAGATGTC 1200
DB 1329 AGATCTGGCACTTCTATATAGTCACTCAAGCAAAAACATATTAACATGGAAGATGTC 1388
QY 1201 TCAGCAAAATGCTTCACTGCTCCAAAAGATCAAGAGAGGTGAATGGAAGAGGTAC 1260
DB 1389 TCAGCAAAATGCTTCACTGCTCCAAAAGATCAAGAGAGGTGAATGGAAGAGGTAC 1448
QY 1261 TCACCACAGACAAATGCAATTTTAACTTCTTAAAGAAAGACATCCGATAGT 1320
DB 1449 TCACCACAGACAAATGCAATTTTAACTTCTTAAAGAAAGACATCCGATAGT 1508
QY 1321 TCTGATCTTTTGAAGAGCTGTATAACAATCAAGCACTTCCAACTCTTGTCCAGGA 1380
DB 1509 TCTGATCTTTTGAAGAGCTGTATAACAATCAAGCACTTCCAACTCTTGTCCAGGA 1568
QY 1381 AAAAATCTTTTCAATTTGAGAGACCCGACACTCGAATGGAACCTGTACCAAGTGTTT 1440
DB 1569 AAAAATCTTTTCAATTTGAGAGACCCGACACTCGAATGGAACCTGTACCAAGTGTTT 1628
QY 1441 GGGAAATCTGCAAAATGCTATTTAAGAAAACTACTGAATATGACAGATCAGGCCA 1500
DB 1629 GGGAAATCTGCAAAATGCTATTTAAGAAAACTACTGAATATGACAGATCAGGCCA 1668
QY 1501 AACCGGAGCTTCCAGGGCAATCCAGATTTGCAAGAGACACATCAAAAAATGCTGGACT 1560
DB 1689 AACCGGAGCTTCCAGGGCAATCCAGATTTGCAAGAGACACATCAAAAAATGCTGGACT 1748
QY 1561 GATACAAAAAGTCAAAAAAGACCTGAGTCTGTATTAATGCACTCTGTATTAACAGCAA 1620
DB 1749 GATACAAAAAGTCAAAAAAGACCTGAGTCTGTATTAATGCACTCTGTATTAACAGCAA 1808

QY 1621 AATACCATGAATATATGACTGCACTTCAAGTAAACCTGAGATATCCAAAGAAATGT 1680
DB 1809 AATACCATGAATATATGACTGCACTTCAAGTAAACCTGAGATATCCAAAGAAATGT 1868
QY 1681 GTTTTGGCTGCAATCTCTTTCTGAAACAGACAGACTAGGGGTATGAGGCCACATGG 1740
DB 1869 GTTTTGGCTGCAATCTCTTTCTGAAACAGACAGACTAGGGGTATGAGGCCACATGG 1928
QY 1741 GGTATCAGAAATCGTACATTTAAGAAAGCATTAATCCGTTGGTGTCCAGGTTAAA 1800
DB 1929 GGTATCAGAAATCGTACATTTAAGAAAGCATTAATCCGTTGGTGTCCAGGTTAAA 1988
QY 1801 CCAATCAGCAGAAAAACAAAAGGCTGTGTGAGCAATCTTGATTCAGAGAGGTGTG 1860
DB 1989 CCAATCAGCAGAAAAACAAAAGGCTGTGTGAGCAATCTTGATTCAGAGAGGTGTG 2048
QY 1861 GTGAGCTTGTAAGAGATGATCTCAAGAAATATGGAAGAAAGTTCTTCAGATATCT 1920
DB 2049 GTGAGCTTGTAAGAGATGATCTCAAGAAATATGGAAGAAAGTTCTTCAGATATCT 2108
QY 1921 AGTATGGAATACGATCACTATTTATATCAAAATGGGTAGAGGTTTCTCTTGCT 1980
DB 2109 AGTATGGAATACGATCACTATTTATATCAAAATGGGTAGAGGTTTCTCTTGCT 2168
QY 1981 GATAGACACCCCTCACCTTGAACAATCACTGATAGTACAGCTTTGACAAATTTACAGAA 2040
DB 2169 GATAGACACCCCTCACCTTGAACAATCACTGATAGTACAGCTTTGACAAATTTACAGAA 2228
QY 2041 AAATCTGGCAAAATATCAATATGCTTCCAGTTTGTGACAGTTGTATGATCTAAATCT 2100
DB 2229 AAATCTGGCAAAATATCAATATGCTTCCAGTTTGTGACAGTTGTATGATCTAAATCT 2288
QY 2101 CCCAAATCACTTATTTTAAAGATATGCTAATGCAATTTTATGAGAAATTCCTCTGT 2160
DB 2289 CCCAAATCACTTATTTTAAAGATATGCTAATGCAATTTTATGAGAAATTCCTCTGT 2248
QY 2161 GCTGATTTTGAAGTTTGATGATGGGTAAATATCAACAAAACGAAGATTTCAAT 2220
DB 2349 GCTGATTTTGAAGTTTGATGATGGGTAAATATCAACAAAACGAAGATTTCAAT 2408
QY 2221 CAGGTATTTGAAAGACAGGGAAGCTTACACTTTAAAGTGAAGTAAATAGC 2280
DB 2409 CAGGTATTTGAAAGACAGGGAAGCTTACACTTTAAAGTGAAGTAAATAGC 2468
QY 2281 TTGAAAGAGGATTAATAATGTAATGACACATGCTAATAGGCTCATGCTATTTGTTA 2340
DB 2469 TTGAAAGAGGATTAATAATGTAATGACACATGCTAATAGGCTCATGCTATTTGTTA 2528
QY 2341 GCACTGGAATCAATTAATTTGAAAGAGAAAGAAAGTAAAGAGTCTCCCTTTTCCCA 2400
DB 2529 GCACTGGAATCAATTAATTTGAAAGAGAAAGAAAGTAAAGAGTCTCCCTTTTCCCA 2588
QY 2401 ATATATCATGGAAGAAACCTGTATGATTAATGCTAAGGCTTATCACTCTCTCT 2460
DB 2589 ATATATCATGGAAGAAACCTGTATGATTAATGCTAAGGCTTATCACTCTCTCT 2648
QY 2461 TCTGTGATTCAAATTAACCAACGAGATGAGCACTTTTCAACGAATGTCAATGAT 2520
DB 2649 TCTGTGATTCAAATTAACCAACGAGATGAGCACTTTTCAACGAATGTCAATGAT 2708
QY 2521 AGTGTGCTTCTTCAACAGGACCAATCTTAAATCCCTCTATAGTTTCAAAATGAGA 2580
DB 2709 AGTGTGCTTCTTCAACAGGACCAATCTTAAATCCCTCTATAGTTTCAAAATGAGA 2768
QY 2581 CTGTGCTTAAACAATCACTGAGCTTGAACAGCACTCTCTTAATATGCTTAAAGATGT 2640
DB 2769 CTGTGCTTAAACAATCACTGAGCTTGAACAGCACTCTCTTAATATGCTTAAAGATGT 2828
QY 2641 CTTCCTAAATCAGCAAACTTTTGAATCTGTTTGTGAAAAATGTTGTTGGCTCA 2700
DB 2829 CTTCCTAAATCAGCAAACTTTTGAATCTGTTTGTGAAAAATGTTGTTGGCTCA 2888
QY 2701 CAGTTAACTAGTGAAGCTGTGTGGTTCAATTAATGATGGGTCCCAAGTGTGTGAG 2760
```

|||||  
Db 2889 CAGTTAACTAGTGGAGCTGTGGGTTCACTTTAAAGATGGGCCAGTTGGTTGGTGGCAG 2948  
Qy 2761 GCAAGAGTCTCTTCTATCACTTATACCTCAACCAATGGTCAAAACAATGGATGGAGAA 2820  
Db 2949 GCAAGAGTCTCTTCTATCACTTATACCTCAACCAATGGTCAAAACAATGGATGGAGAA 3008  
Qy 2821 AATGAAATTTACGACCTACATCAACGAATTAACAGTGTCTGCTTCAATCCCTTTG 2880  
Db 3009 AATGAAATTTACGACCTACATCAACGAATTAACAGTGTCTGCTTCAATCCCTTTG 3068  
Qy 2881 ATGTTTCTATCCGACCTCTATTTTCATTTGA 2913  
Db 3069 ATGTTTCTATCCGACCTCTATTTTCATTTGA 3101  
RESULT 5  
AR083569 3447 bp DNA linear PAT 01-SEP-2000  
LOCUS Sequence 3 from patent US 5976893.  
ACCESSION AR083569  
VERSION AR083569.1 GI:10010342  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3447)  
AUTHORS Dennis,J.W., Heffernan,M. and Fode,C.  
TITLE Methods for identifying binding partners, agonists, and antagonist  
JOURNAL of a serine/threonine tyrosine kinase  
FEATURES Patent: US 5976893-A 3 02-NOV-1999;  
source Location/Qualifiers  
1..3447  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match 64.5%; Score 1879; DB 6; Length 3447;  
Best Local Similarity 80.9%; Pred.No. 0;  
Matches 2362; Conservative 0; Mismatches 410; Indels 147; Gaps 9;

Qy 1 ATGGCAGCTCGATCGGGAGAAATCGAGATTTTAAAGTGAATCTGCTGTGA 60  
Db 206 ATGGCGCGCTCGATCGGGAGAGATCGAGACTTTAAAGTTGAATCTACTCGTTAA 265  
Qy 61 GGATCATTTGCTGTGTCTTACAGAGCTGAGTCCATTCACACCTGTTTGAAGTTGCAATC 120  
Db 266 GGATCATTTGCTGTGTCTTACAGAGCTGAGTCCATTCACACCTGTTTGAAGTTGCAATC 325  
Qy 121 AAAATGATGATGAAGAAAGCCATGTCAGAGAGAAATGGTACAGAGATCCAAATGAG 180  
Db 326 AAAATGATGATGAAGAAAGCCATGTCAGAGAGAAATGGTACAGAGATCCAAATGAG 385  
Qy 181 GTGAAATATCATTTGCCAATTTGAACATCTTCTATCTTGGAGCTTATATCTTTTGA 240  
Db 386 GTGAAATATCATTTGCCAATTTGAACATCTTCTTGGAGCTTATATATTTTGA 445  
Qy 241 GATAGCAATTTATGTATCTGTATTTGAATGTCCATTAATGAGAAATGAACAGTAT 300  
Db 446 GATAGCAATTTATGTATCTGTATTTGAATGTCCATTAATGAGAAATGAACAGTAT 505  
Qy 301 CTAAAGAAATGAGTGAACCCCTTCTGAGAAATGAGAGTCCGACCTTCAATGACCGAGATC 360  
Db 506 CTAAAGAAATGAGTGAAGCCCTTCTGAGAAATGAGAGTCCGACCTTCAATGACCGAGAT 565  
Qy 361 ATCACAGGAGTGTGTATCTTCAATCTCAATGTATCTACACCGGAGCTCAACATTTCT 420  
Db 566 ATCACAGGAGTGTGTATCTTCAATCTCAATGTATCTACACCGGAGCTCAACATTTCT 625  
Qy 421 AACCTCTACTGACTCTGTATATGAACATCAAGATTGCTGATTTGGGCTGGCACTCAA 480  
Db 626 AACCTCTACTGACTCTGTATATGAACATCAAGATTGCTGATTTGGGCTGGCAAGCGAG 685

Qy 481 CTGAAATATGCACATGAAAGACCTATACATTTATGTGAATCTCTACTATCAATTTCCACA 540  
Db 686 TTGAATATCCCACTGAAATGAAGACCTATACATCTGTGGGACCTCTATATTTATCCACA 745  
Qy 541 GAAATTCACATCGAAGTGCACATGGCTTGAATCTGATGTTTGGTCCCTGGGCTGTATG 600  
Db 746 GAAATTCACATCGAAGTGCACATGGACTGATGATCTGATTTTGGTCAATTTGGGCTGTATG 805  
Qy 601 TTTTATACATTAATTATGGGAGACCCCTCGACATGACAGACAGTCAAGAACATTA 660  
Db 806 TCTTATACATTAATTATGGAGACCCCTTGTGACATGACAGACATCAAGAACATTA 865  
Qy 661 AATTAAGTATTTGGCAGATTTATGAATGCCATCTTTTGTCAATAGAGCCAGAGAC 720  
Db 866 AACAAAGTATCTTGGCAGATTTATGAATGCCATCTTTTGTCAAGAGGCCAGAGAC 925  
Qy 721 CTATATCACAGTTACTTGTGAGAAATCCAGACAGATCGTTTAAATGTCTTCAATTTG 780  
Db 926 CTATATCACAGTTACTTGTGAGAAATCCAGACAGATCGGTAAATGTCTTCTGTGTTG 985  
Qy 781 GACATCTCTTATATGCCCCGAAATCTTCAAGAAATGAATTTAGGAATGTGGA 840  
Db 986 GACATCTCTTATATGCAAGAAATCTTCAAGAAATGAATTTAGGAATGTGGA 1045  
Qy 841 GACTCAATGATAGTGGGATGCCCAATTTCTACTGCAATTTACAGCTTCTTCAAGTACC 900  
Db 1046 GACTCAATGATAGTGGGATGCCCAATTTCTACTGCAATTTACAGCTTCTTCTGATAC 1105  
Qy 901 AGTATAGTGTAGTTTATTTGACAAAGAGACCTTTGATTTGTGACCACTCCCAAT 960  
Db 1106 AGTATAGTGTAGTTTATTTGACAAAGAGACCTTTGATTTGTGACCACTCCCAAT 1162  
Qy 961 AAAATGACCTGTATTTCCAAAGAAATGAATGAATGATTTTCTTCTCAGAGATGGA 1020  
Db 1163 AAAATGACCTGTATTTCCAAAGAAATGAATGAATGATTTTCTTCTCAGAGATGGA 1219  
Qy 1021 AACAGTTTATCTACATGAGGGAAT-----CAAGAAACAGTATAGTGAAGGGA 1074  
Db 1220 AATATTTTATCTACATGAGGGAATCCAGAAACAGTATAGTGAAGGGA 1279  
Qy 1075 AAGATATTTCAAGATGCAAGAAAGGCCAATTCATGATCTTCTGATAGCTTATTC 1134  
Db 1280 AAGATATTTCAAGATGCAAGAAAGGCCAATTCATGATCTTCTGATAGCTTATTC 1339  
Qy 1135 TCTGATAGTCTGCACTTCTAATAGTCTTCAAGCAAAACATATCAATGAGACGA 1194  
Db 1340 TCTGATAGTCTGCAAGCCCTCTAA---TCAGTCTGAGCAAAACATATCAGTGAACGT 1396  
Qy 1195 TGTCACTCAGCAAAATGCTTCAAGTCTCAAAAGATCAGAGAGAGTGAATGAAGAG 1254  
Db 1397 TGTCACTCAGTGAATGCTTCAAAAGCTTGAAGATC----- 1434  
Qy 1255 AGGTACTCACCAAGACAAATGCAATTTTAACTTTTAAAGAAAAACATCC 1314  
Db 1435 ----- 1434  
Qy 1315 AGTATGTTGATCTTTTGAAGACCTGATTAACATCAAGACCTCCCATGATCTTTGT 1374  
Db 1435 -----ACTGATGAAATTCACACAGTTCCATATCATTTGT 1471  
Qy 1375 CCAGGAAAAATCTCTTTTCAATTTGAGAGCCGAGACCTCAGACTGAACCTGACAG 1434  
Db 1472 CTAGGAAAAATCTCTTTTCAATTTGAGAGCCGAGACCTCAGACTGAATGATGAGCAG 1531  
Qy 1435 TGGTTTGGGAATCTGCAATTAATGCTCTTTTAAAGAAAACTACTGAATATGACAGATC 1494  
Db 1532 TGGTTTGGGAATCTGCAATTAATGCTCTTTTAAAGAAAACTACTGAATATGACAGATC 1591  
Qy 1495 AGCCCAAAACCGGAGCTTCCAGAGCCATCCAGATTTTGGAGAAAGACACATCAAAAATGCGC 1554  
Db 1592 AGCCCAAAACAGAGATTTTCCAGAGCTATCCAGATTTTGC---AGACACGTTTGAAGAACGT 1648  
Qy 1555 TGAATGATCAAAAGTCAAAAGAACTGTGATGCTTGTGATATGACATTTCTGTAA 1614

Db	2709	----	GTCTTCCTTAATCTGCACAGCTTTTGAATCTGTTTGTGAAAAATGTTGGTGG	2764
Qy	2695	GCTACACAGTTAACTAGTGTGAGACTGTGTGGGTTCAATTAAATGATGGGTCACAGTTGGT	2754	
Db	2765	GCACACAGCTTAACTAAGGAGAGCTGTGTGGGTTCAATTAAATGATGGGTCACAGTTGGT	2824	
Qy	2755	GTGACAGGACGAGTGTCTTATCAGTTATACCTCCACCAATGGTCAACCAATAGTAT	2814	
Db	2825	GTTCACAGGAGAGTATCTTTCATCATGTTACATACACAGATGGTCAGAACACTAGTAT	2884	
Qy	2815	GGAGAAAATGAAAAATTAACAGACTACATCAACAGAAATTAACAGTGTCTGTCTTCATC	2874	
Db	2885	GGAGAAAATGAAAAATTAACCTGAATATCATCAACAGAAATTAACAGTGTCTTTCATC	2944	
Qy	2875	CTTTGATGTGTTTCTTAATCGAGCTCTTAATTTTCATGTA	2913	
Db	2945	CTTTCGATGTTTCTTAATTCACATCTCTTAATTTTCATGTA	2983	
RESULT 6				
LOCUS	156871		3447 bp	DNA
DEFINITION	Sequence 3 from patent US 5650501.			linear
ACCESSION	156871			
VERSION	156871.1		GI:2477284	
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 3447)			
AUTHORS	Dennis,J.W., Heffernan,M. and Fode,C.			
TITLE	Serine/threonine kinase and nucleic acids encoding same			
JOURNAL	Patent: US 5650501-A 3 22-JUL-1997;			
FEATURES	Location/Qualifiers			
source	1..3447			
ORIGIN				
	/organism="unknown"			
	/mol_type="unassigned DNA"			
Query Match	64.5%;	Score 1879;	DB 6;	Length 3447;
Best Local Similarity	80.3%;	Pred. No. 0;		
Matches 2362;	Conservative	0;	Mismatches 410;	Indels 147;
			Gaps	9;
Qy	1	ATGCGCACCTGCATCGGGGAGAGATCGAGGATTTTAAAGTTGGAAATCTGCTGGTAA	60	
Db	206	ATGCGCGCTGCATCGGGGAGAGATCGAGACTTTAAAGTTGAAATCTCGTAA	265	
Qy	61	GGATCATTTGCTGTGTCTTACAGAGCTGAGTCAATCACTGTTGGAAATGCAATC	120	
Db	266	GGATCATTTGCTGTGTCTTACAGAGCTGAGTCAATCACTGTTGGAAATGCAATC	325	
Qy	121	AAATGATATGATGAAGAACCATGTACAAAGAGAGAAATGTACAGAGTCCAAATGAG	180	
Db	326	AAATGATATGATGAAGAACCATGTACAAAGCTGGAATGTACAGAGTCCAAATGAG	385	
Qy	181	GTGAATAATCATTTGCCAATGTAAACATCTTATCTTGAAGCTTATATCTATTTTGA	240	
Db	386	GTGAATAATCATTTGCCAATGTAAACATCTTATCTTGAAGCTTATATCTATTTTGA	445	
Qy	241	GATAGCAATTAATGTGATCTGTATTAATAATGTGCCATATGAGAAATGAACAGTAT	300	
Db	446	GATAGCAATTAATGTCTACTGTATTTGAATGTGCCCAATGAGAAATGAACAGTAT	505	
Qy	301	CTAAAGATATGAGTAAACCTTCTCAGAAATATGAAGCTGCACCTTCATGCACAGATC	360	
Db	506	CTGAAGATATGAGTAAAGCTTCTCAGAAAGGAGCTTATGAGCACTTCATGCACAGATC	565	
Qy	361	ATACAGGAGATGTGTATCTTCTATCTCATGTATATCAACCGGAACTTCACCTTCT	420	
Db	566	ATACAGGAGATGTATATCTTCTATCTCATGTATATTCACCGGAACTTCACCTTCT	625	
Qy	421	AACTCTTACTGACCTGATATATGAACATCAAGATGCTGATTTTGGGCTGGCAACTCA	480	

DB 626 AACATCTTACTTACGCGGAATATGAACATTAATAATGTGACTTGGACTAGCAACGCG 685  
QY 481 CTGAAAATCCCATGATAAAGCACTATATTAATGTGAACTCCCTAATCTACATTTTCCCA 540  
DB 686 TTGAATATCCCATGATAAAGCACTATATTAATGTGAACTCCCTAATCTACATTTTCCCA 745  
QY 541 GAAATTCGCACTGAAAGTGCACATGCGCTTGAATCTGATGTTGTCCTGGGCTGTATG 600  
DB 746 GAAATTCGCACTGAAAGTGCACATGCGCTTGAATCTGATGTTGTCCTGGGCTGTATG 805  
QY 601 TTTTATATCAATTAATTCGCGAGACCACTTCCTGACATGACAGTCAAGAACATTA 660  
DB 806 TCTTATACCTTACTTATGTGAAGACCACTTTTGAACATGACACAGTCAAGAACATTA 865  
QY 661 AATAAAGTATATGCGACATTAATGAAGCCATCTTTTGTGATATGAGCGCAAGAC 720  
DB 866 AACAAAGTATCTGCGACATTAATGAAGCCATCTTTTGTGACAGAGCGCCAGAC 925  
QY 721 CTTAATCACCAATTAATTCGTAGAAATTCAGCAGATCGTTAATGCTGCTTCAATTTG 780  
DB 926 CTTATCCACAGTACTTGTGAGAAACCTTCGACGATCGGTAAATGCTGCTTCTGTGTTG 985  
QY 781 GACCATCTTTTATATGTCGGAATTCCTCAACAAAAGTAAAGATTTAGGAACTGTGAA 840  
DB 986 GACCATCTTTTATATGTCGGAATTCCTCAACAAAAGTAAAGATTTAGGAACTGTGAA 1045  
QY 841 GACTCAATGATATGAGGCGATGCGCAATTTCTACGCAATTAACACTTCTTCCATATCC 900  
DB 1046 GACTCAATGATATGAGGCGATGCGCAATTTCTACGCAATTAACACTTCTTCTGTATCC 1105  
QY 901 AGTATATGATGATGATTTATTTGACAAAAGAGACTTTTGAATGGTCAAGCCATCCCAAT 960  
DB 1106 AGTTGAGTGCAGCTTACTTGAC---AGAGACTTTTGTGTTGTTGTTGTTGTTGTTGTTG 1162  
QY 961 AAAATGACTGATTTTCCAAAGAAATTAAGTCAACTGATTTTCTTCTTCAAGAGATGA 1020  
DB 1163 AAAATGACTGATTTTCCAAAGAAATTAAGTCAACTGATTTTCTTCTTCAAGAGATGA 1219  
QY 1021 AACAGTTTATATCTCATGTTGGGGAAT-----CAAGAAACCGTAAATGTGAAGGGA 1074  
DB 1220 AATTAATTTTGTATCTCAATGAGGGAATTCAGAACAGAAAGCTAAATGTAGGGAACG 1279  
QY 1075 AGAGTATTCAGAGTSCAGAGAAAGGCGCAATTCGATACCTTGTAGAGCTTATTC 1134  
DB 1280 AGAGTATTCAGAGTSCAGAGAAAGGCGCGCATTTCTGATACCTTGTAGAGCTTATTC 1339  
QY 1135 TCTGATAGATCTGGCACTTCTATATAGTCAATCTCAAGCAAAAACATATCAATGAACGA 1194  
DB 1340 TCTGATAGAGCCAGCCCTCTTA---TCAGTCTCGAGCAAAAACATATCAATGAACGA 1396  
QY 1195 TGTCACTCAGCAAAATGCTTTCAGTGTCCAAAGATCAAGAGAGGTGAAAATGAAGAG 1254  
DB 1397 TGTCACTCAGTGAATGCTTTCAGAGCCTAGAGATC----- 1434  
QY 1255 AGATACCTCAACCAAGACAAATGCGCAATTTTAACTTTTAAAGAAAAGACATCC 1314  
DB 1435 ----- 1434  
QY 1315 AGTATGTTCTGATCTTTTGAAGACTGATTAACATCAAGCACTCTCAATCATCTTTGT 1374  
DB 1435 -----ACTGATGAAATTCACACAGTTCCATCATCATTTGT 1471  
QY 1375 CCAAGAAAAATCTCTTTTTCATTTGCAAGACCGCACTTCAGATGAAACCGTAAACAG 1434  
DB 1472 CTAAGAAAAATCTCTTTTTCATTTGCAAGACCGCACTTCAGATGAAATGGAACAGAG 1531  
QY 1435 TGGTTTGGGAATCTGCAAAATTAATGCTCAATTAAGAAAACTAATGATATGACAGATC 1494  
DB 1532 TGGTTTGGGAATCTGCAAAATTAATGCTCAATTAAGAAAACTAATGATATGACAGATC 1591  
QY 1495 AGCCCAACCGGAGCTTCAAGGCGCATTCAGATTTTCAAGAGCAATCAAAAATGCGC 1554  
DB 1592 AGCCCAACCGGAGCTTCAAGGCGCATTCAGATTTTCAAGAGCAATCAAAAATGCGC 1648

QY 1555 TGAAGTATACAAAGTCAAAAAGAACTGTAGTCTTGTATATGACATTCGTAAAA 1614  
DB 1649 TGAAGTATACAAAGTCAAAAAGAACTGTAGTCTTGTATATGACATTCGTAAAA 1708  
QY 1615 CAGCAAAATACATGAAATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAA 1674  
DB 1709 CAGTGAAGTCCATGAAATATATGAGTGCATATCAATTAAGCTGAGTCAAGCAG 1768  
QY 1675 GAATGTTTGGCTGCAATCTCTTCTTGAACAGAGCAAGACTAGGGGTATGAGCCA 1734  
DB 1769 GA-----GCGGGCTCAATCTCAATCTGAACAAAGCAAGAAATGAAGTGAAGTG 1822  
QY 1735 CCATGGGTTTATCAGAAATGATCAATTAAGAGACATTAATCTCCGTTGGTGTCCACAG 1794  
DB 1823 AACTGAGTTTACAGAAACCTTAAAGAAATTAATCACTCTCTGATGTTCTACAG 1882  
QY 1795 TTAACCAATCAGACAGAAACCAAAAAGGCTGTGTAGACATCTTGAATCAAGAGAG 1854  
DB 1883 TTAAGCCAAATCAGACAGAAACCAAAAAGGCTGTGTAGACATCTTGAATCAAGAGAG 1942  
QY 1855 GTGTGTGTGAGCTTGTAAAGAGTATGATCTCAAGAAATATGTAAAGAAAGTTCTCAG 1914  
DB 1943 GTGTGTGTGAGCTTGTAAAGAGTATGATCTCAAGAAATATGTAAAGAAAGTTCTCAG 2002  
QY 1915 ATATCTAGTATGAAATATGATCACTATTTATTAATCCAAATGTGTAGAGTTTCT 1974  
DB 2003 ATATCTAGTATGAAATATGATCACTATTTATTAATCCAAATGTGTAGAGTTTCT 2062  
QY 1975 CTGTGTATGACCAACCTCCTACCTACTGACCAATCAGATGAGTACAGCTTTGACATTTA 2034  
DB 2063 CTGTGTATGACCAACCTCCTCCTACTGACCAATCAGATGAGTACAGCTTTGACATTTA 2122  
QY 2035 CCAAGAAATATCTGCGGAAATATCAATATGCTTCCAGTTGTATACAGTTGTAGATCT 2094  
DB 2123 CCAAGAAATATCTGCGGAAATATCAATATGCTTCCAGTTGTATACAGTTGTAGATCT 2182  
QY 2095 AAATCTCCCAAAATCACTATTTTCAAGATATGCTAAATGCAATTTTGAATGAAATCT 2154  
DB 2183 AAATCTCCCAAAATCACTATTTTCAAGATATGCTAAATGCAATTTTGAATGAAATCT 2242  
QY 2155 CCTGTGTGATTTTGAAGTTTGTGTTTATGATGAGGTTAAATAACAAACAGAAAGT 2214  
DB 2243 CCTGTGTGATTTTGAAGTTTGTGTTTATGATGAGGTTAAATAACAAACAGAAAGT 2302  
QY 2215 TTCAATCAGTATGTAAGAAAGCAGGAACTTCACTTTAAAGTAAAGTGAAGT 2274  
DB 2303 TTCAATCAGTATGTAAGAAAGCAGGAACTTCACTTTAAAGTAAAGTGAAGT 2362  
QY 2275 AATAGCTGAAGAGAGATTAATAATGTAATGAGCAATGATGAGGTTCAATGAT 2334  
DB 2363 AATAGCTGAAGAGAGATTAATAATGTAATGAGCAATGATGAGGTTCAATGAT 2422  
QY 2335 TGTATGACATGGAATTCATATTTCAAGAAAGAAAGAACTAGAGTCTCCCTT 2394  
DB 2423 TGTATGACATGGAATTCATATTTCAAGAAAGAAAGAACTAGAGTCTCCCTT 2482  
QY 2395 TTCCCAATATCAATGAAAGAAACCTGTGATGATGATTCATCTTCAAGGCTTATCACT 2454  
DB 2483 TTCCCAATATCAATGAAAGAAACCTGTGATGATGATTCATCTTCAAGGCTTATCACT 2542  
QY 2455 CCTCTCTGTGGAATTCATATTAACCAACGAGAGATGAGATCTTTCAAGAAATGCT 2514  
DB 2543 CCTCTCTGTGGAATTCATATTAACCAACGAGAGATGAGATCTTTCAAGAAATGCT 2599  
QY 2515 ATGATAGTGTGCTTCTTCCACAGAGCCCAATCTTAACTCCCTATGTTTCAAT 2574  
DB 2600 GTGAATATGAGGCTTCTTCCACAGAGCCCAATCTTAACTCCCTATGTTTCAAT 2659  
QY 2575 GAAAGCTTGTGTTTCAATCACTACAGTTTGTGAACAGACATCTTCTTAAATGCTTAA 2634  
DB 2660 GAAAGCTTGTGTTTCAATCACTACAGTTTGTGAACAGAGGCTTCTTCAAT----- 2708

Qy 2635 GATTGCTCTTCCAAATCAGACAACCTTTGAAATCTGTTTTGAAAAAATGTTGGTGG 2634  
Db 2709 ----GTCTTCTTAATCTGCACAGCTTTTGAAATCTGTTTTGAAAAAATGTTGG 2764  
Qy 2635 GCTACACAGTTAACTAGTGAAGCTGTGTGGGTTCAGTTTAATGATGGGTCCAGTTGGTT 2754  
Db 2765 GCTACACAGCTAACTAGTGAAGCTGTGTGGGTTCAGTTTAATGATGGGTTCAGTTGGTT 2824  
Qy 2755 GTGAGGAGGAGAGTGTCTTCTATCAGTTAATCCCAACCAATGCTCAACAGTAT 2814  
Db 2825 GTCAGGAGGAGAGTGTCTTCTATCAGTTAATCCCAACCAATGCTCAACAGTAT 2884  
Qy 2815 GAGAAAAATGAAAAATTAACAGACTACATCAACAGAAATTAACAGTGTCTTCTTCATC 2874  
Db 2885 GAGAAAAATGAAAAATTAACAGACTACATCAACAGAAATTAACAGTGTCTTCTTCATC 2944  
Qy 2875 CTTTGTGATGTTTCTTAATCCGACTCTCTTAATTTTCATGTA 2913  
Db 2945 CTTTGTGATGTTTCTTAATCCGACTCTCTTAATTTTCATGTA 2983

RESULT 7  
AX306014 3447 bp DNA linear PAT 11-DEC-2001  
LOCUS Sequence 765 from Patent WO018188.  
DEFINITION AX306014  
ACCESSION AX306014 GI:17645355  
VERSION AX306014.1  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Bukayocia; Metzoz; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T. and Ishii, Y.  
AUTHORS Method for examining ischemic conditions  
TITLE Patent: WO 018188-A 765 22-NOV-2001;  
JOURNAL School Juridical Person Nihon University (JP)

FEATURES  
source location/Qualifiers  
1..3447

ORIGIN  
/organism="Mus musculus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:10090"

Query Match 64.5%; Score 1879; DB 6; Length 3447;  
Best Local Similarity 80.9%; Pred. No. 0;  
Matches 2362; Conservative 0; Mismatches 410; Indels 147; Gaps 9;

Qy 1 ATGGCGACCTGCATCGGGGAGAAATCGAGATTTTAAAGTTGAAATCTGCTGTAA 60  
Db 206 ATGGCGGCGTGCATCGGGGAGAGATCGAGACTTTAAAGTTGAAATCTGCTGTAA 265  
Qy 61 GGATCAATTTGCTGTGTCTACAGAGCTGATCACTGATTTGAAAGTTGCAATC 120  
Db 266 GGATCAATTTGCTGTGTCTACAGAGCTGATCACTGATTTGAAAGTTGCAATC 325  
Qy 121 AAAATGATGATGAAAGGCAATGTAACAAGGAGAAATGTAACAAGAGTCCAAATGAG 180  
Db 326 AAAATGATGATGAAAGGCAATGTAACAAGGAGAAATGTAACAAGAGTCCAAATGAG 385  
Qy 181 GTGAAATATACATTTGCAATGTAACAATCTTATCTTGAGACTTTAATCTAATTTGA 240  
Db 386 GTGAAATATACATTTGCAATGTAACAATCTTATCTTGAGACTTTAATCTAATTTGA 445  
Qy 241 GATGCAATTTATGCTATCTGTATTAAGAAATGTCCTAATGAGAAATGAAAGATAT 300  
Db 446 GATGCAATTTATGCTATCTGTATTAAGAAATGTCCTAATGAGAAATGAAAGATAT 505  
Qy 301 CTAAAGATAGAGTGAAGACCTTCTCAGAAATGAAAGTTCGACCTTCATGCAACAATC 360  
Db 506 CTGAAGAAACAGAAATGAAGCTTTCTCAGAAAGGAGAGTTCGACCTTCATGCAACAATC 565  
Qy 361 ATCAGAGGAGATGTTGATCTTCAATTCATGATTAACAACCGGAGACTCACAATTTCT 420

Db 566 ATCAGAGGAGATGTTGATCTTCAATTCATGATTAACAACCGGAGACTCACAATTTCT 625  
Qy 421 AACCTCTTCTAGCTCTGTAATATGAACATCAAGATTTGCTGATTTTGGGCTGCAACTCA 480  
Db 626 AACATTTTCTTACGGGGAATATGAACATTAATAATTTGCTGACTTTGACTACCAACGAG 685  
Qy 481 CTGAAATGCAATGAAAGGCACTATACATATGATGAGAACTCTTAATTTTCAACA 540  
Db 686 TTGAATATGCAATGAAAGGCACTATACATCTGTGGGACTCTTAATTTTCAACA 745  
Qy 541 GAAATGCACTGAAAGTGCATATGCTCTGATATGATTTGGTCTGCTGGCTGATG 600  
Db 746 GAAATGCACTGAAAGTGCATATGATCTGATATGATTTGGTCTGCTGGCTGATG 805  
Qy 601 TTTTATACATTTACTATGAGGAGAACCCCTTCGACATGACAGAGTAAGAACATTA 660  
Db 806 TCTTATACGTTATCTTATGAAAGACCTTTTGAACCTGACAGTCAAGAACACATG 865  
Qy 661 AATTAAGTATGTTGGCAGATTTATGAATGCAATCTTTTGTCAATGAGGCAAGAC 720  
Db 866 AACAAAGTATGCTCTGGCAGATTTATGAATGCAAGCTTTTGTCAAGAGGCCAAGAC 925  
Qy 721 CTTATTCACAGTTACTTGTGAAATTCAGACATGCTTTAAGTCTGCTTCAGATTTG 780  
Db 926 CTTATTCACAGTTACTTGTGAAATTCAGACATGCTTTAAGTCTGCTTCAGATTTG 985  
Qy 781 GACCATCTTTTATGCTCCGAATTTCTCAACAAAAGTAAAGATTAAGAACTGTGAA 840  
Db 986 GACCATCTTTTATGCTCCGAATTTCTCAACAAAAGTAAAGATTAAGAACTGTGAA 905  
Qy 841 GACTCAATTTATGTTGGGCAATTCGCAATTTCTGCAATTAAGCTTCTTCCAGTACC 900  
Db 1046 GACTCAATTTATGTTGGGCAATTCGCAATTTCTGCAATTAAGCTTCTTCCAGTACC 1105  
Qy 901 AGTAAAGTGTATGTTTATTTGCAAAAAGAACTTTTGTGATGCTGACCACTCCAAAT 960  
Db 1106 AGTAAAGTGTGCACTTATTTGCAAAAAGAACTTTTGTGATGCTGACCACTCCAAAT 1162  
Qy 961 AAAATGACTGATTTTCCAAAGAAATTAAGTCAATGATTTTCTTCCAGAGATGA 1020  
Db 1163 AAAATGACTGATTTTCCAAAGAAATTAAGTCAATGATTTTCTTCCAGAGATGA 1219  
Qy 1021 AACATTTTATACAGTGTGGGGAAT-----CAAGAAACAGTAAATATGAAAGGGA 1074  
Db 1220 AACATTTTATACAGTGTGGGGAAT-----CAAGAAACAGTAAATATGAAAGGGA 1279  
Qy 1075 AAGTAAATTCAGATGCAAGAAAGGCAATTTCTGATACCTTGTGAGAGTTATTC 1134  
Db 1280 AAGTAAATTCAGATGCAAGAAAGGCAATTTCTGATACCTTGTGAGAGTTATTC 1339  
Qy 1135 TCTGATGATCTGCACTTCTTAATGTCAGTCTCAAGCAAAAACATTAATGAAAGCA 1194  
Db 1340 TCTGATGATCTGCACTTCTTAATGTCAGTCTCAAGCAAAAACATTAATGAAAGCA 1396  
Qy 1195 TGTCACTCAGCAGAAATGCTTTCAAGTGTCCAAAAGATCGAGAGAGTGAATGAAG 1254  
Db 1397 TGTCACTCAGCAGAAATGCTTTCAAGTGTCCAAAAGATCGAGAGAGTGAATGAAG 1434  
Qy 1255 AAGTATCACCCACAGACAAATGCAACATTTTAACTTTTAAAGAAAGACATCC 1314  
Db 1435 ----- 1434  
Qy 1315 AGTAAATTCAGATCTTTGAAAGACCTGATTAACATCAAGCACTCTCAATCATTTGT 1374  
Db 1435 -----ACTGATGAAATTAACACAGATTTCAATCATTTGT 1471  
Qy 1375 CCGAGAAAAATCTCTTTTCATTTGCAAGCCGACACTCAGATGAACCTGAACAAG 1434  
Db 1472 CTAGGAAAAATCTCTTTTCATTTGCAAGCCGACACTCAGATGAATGTTAACAAG 1531  
Qy 1435 TGGTTTGGGAATCTGCAATTAATGCTCAATTTAAGAAAACTACGATATGACAGCATC 1494



Db 1532 TGGTTGGGAATCTGCAATGAATGCTCATTTAGAGAAACTAATGAGCACCAACCGTT 1591  
 Qy 1455 AGCCCAACCGGAGCTTCCAGGGCCATCCAGATTTGGAGAGAGACATCAAAAATATGCC 1554  
 Db 1592 AGCCCAACGAGATTTCCAGAGCTATCCAGATTTGCG---AGGACCGTTTACGAAACGCT 1648  
 Qy 1555 TGGACTGATACAAAGTCAAAAAGAACTGTGATGCTTGTGATATGCAATTCGTGAAA 1614  
 Db 1649 TGGATGACACAGAGAGCCAGAGAAATGCTGATCTTCTGCAATGTTTATGCTGTGAAG 1708  
 Qy 1615 CAGCAAAATACAGTAATATATATGACTGCACTTCAAGTAACTTGAGATTAATCCAA 1674  
 Db 1709 CAGCTGAGTGCCATGAAATATACATGAGTGCAATACATAGCTTGAAGGCTCATGCCA 1768  
 Qy 1675 GAATGTGTTTGGCTCAGATCTCTCTTTTGAACAGAGCAAGATAGGGGTATGAGCC 1734  
 Db 1769 GA-----GCCGGGGCTCATCTCTCAATCTGAAACAAAGCAAAATGAAAGTATGAGTGG 1822  
 Qy 1735 CAGTGGGGTTATCAGATCTGATCATTTAAGAGCAATTAATCCGTTGGTGTCTCAGAG 1794  
 Db 1823 ACAGTGGGTTACAGAAACCTTACAGAAATTAATTAATCTCTCTGATGCTCAGAGA 1882  
 Qy 1795 TTAACCAATACAGACAGAAACCAAAAGGCTGTGTGAGCATATCTTGATTCAGAGAG 1854  
 Db 1883 TTAACCAATACAGACAGAAACCAAAAGGCTGTGTGAGCATCTTGATTCAGAGAG 1942  
 Qy 1855 GTGTGTGTGAGCTTGTAAAGAGATGATCTCAAGATATGTGAAGAAAGTTCTTCAG 1914  
 Db 1943 GTGTGTGTGAGCTTGTAAAGAGATGATGCTCAAGATATGTGAAGAAAGTTCTTCAG 2002  
 Qy 1915 ATATGATGATGAGAAATACAGATCTATTTATTAATCCAAATGTGTGAGAGTTTCT 1974  
 Db 2003 ATATGATGATGAGAAATACAGATCTATTTATTAATCCAAATGTGTGAGAGTTTCT 2062  
 Qy 1975 CTGTGTGTGAGAAATACAGATCTATTTATTAATCCAAATGTGTGAGAGTTTCT 2034  
 Db 2063 CTGTGTGTGAGAAATACAGATCTATTTATTAATCCAAATGTGTGAGAGTTTCT 2122  
 Qy 2035 CCAGAAATATCTGGCGAAATATCAATATGCTTCCAGGTTGTACAGTTGTAGATCT 2094  
 Db 2123 CCAGAAATATCTGGCGAAATATCAATATGCTTCCAGGTTGTACAGTTGTAGATCT 2182  
 Qy 2095 AAATCTCCAAATACATTTATTTCAAGATATGCTTCAATGATTTGATGAGAAATCT 2154  
 Db 2183 AAATCTCCAAATACATTTATTTCAAGATATGCTTCAATGATTTGATGAGAAATCT 2242  
 Qy 2155 CTTGTGTGATTTTGTAGGTTGTTTATGATGAGGTTTAAATATCACAAACAGAGAT 2214  
 Db 2243 CTTGTGTGATTTTGTAGGTTGTTTATGATGAGGTTTAAATATCACAAACAGAGAT 2302  
 Qy 2215 TTCAATCAGGTGATGAAAGACAGGGAAGCTTACATTTTAAAGTGAAGTGAAGT 2274  
 Db 2303 TTAATTCATATTAATGAGAAACAGGGAATCTTATTAATTTAAATGAAATGATGAGT 2362  
 Qy 2275 AATACCTGAAAGAGAGATTAATAATGATATGAGCAATGATGAGGTCATGATGAT 2334  
 Db 2363 ACCACCTGAAAGAGAGATTAATAATGATATGAGCAATGATGAGGTCATGATGAT 2422  
 Qy 2335 TGTTTAGCACTGGAATCCATTAATTTCAAGAGAGAAAGAAATAGAGAGTCTCCCTT 2394  
 Db 2423 TGTCTGTCACTGGAATCTGTATCTGTGAGAGAAAGAGAGAGAGGCTTCTTCAATTC 2482  
 Qy 2395 TTCCCAATATATAGAGAGAAAGAACTGTGATGATGATGATGATGATGATGATGAT 2454  
 Db 2483 TTCCCAATATATAGAGAGAAAGAACTGTGATGATGATGATGATGATGATGATGAT 2542  
 Qy 2455 CTTCTGTGTGATTTCAATTAATCCAAAGAGAGATGAGCAATTTTCAAGAAATGATG 2514  
 Db 2543 CTTCTGTGTGATTTCAATTAATCCAAAGAGAGATGAGCAATTTTCAAGAAATGATG 2599  
 Qy 2515 ATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2574  
 Db 2600 GTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2659

Qy 2575 GAAGACCTTGTGCTTACAACTACAGCTTCTGGAACAGACATCTTCTTAATAGTCTAAA 2634  
 Db 2660 GAAGACCTTGTGCTTACAACTACAGCTTCTGGAACAGACATCTTCTTAATAGTCTAAA 2708  
 Qy 2635 GATGTCTTCTTAAATACAGCACTTGTGAATGCTTTTGTGAATAATGTGTGTGG 2694  
 Db 2709 ----GCTTCTTAAATACAGCACTTGTGAATGCTTTTGTGAATAATGTGTGTGG 2764  
 Qy 2695 GCTACACAGTTAACTAGTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2754  
 Db 2765 GCTACACAGTTAACTAGTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2824  
 Qy 2755 GTGACGAGAGAGTGTCTTCTTAACTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2814  
 Db 2825 GTGACGAGAGAGTGTCTTCTTAACTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2884  
 Qy 2815 GGAGAAATGAAATAATACAGACTTACATCAACAGAAATTAACAGTGTGTGTGTGT 2874  
 Db 2885 GGAGAAATGAAATAATACAGACTTACATCAACAGAAATTAACAGTGTGTGTGTGT 2944  
 Qy 2875 CTTTGTGATGTTTCTTAATCCAGCTCTTAATTTTCAATGA 2913  
 Db 2945 CTTTGTGATGTTTCTTAATCCAGCTCTTAATTTTCAATGA 2983

RESULT 8  
 MUSA  
 LOCUS MUSA 3447 bp mRNA linear ROD 24-JUN-1994  
 DEFINITION Mus musculus serine/threonine kinase (sak-a) mRNA, complete cds.  
 ACCESSION L29479  
 VERSION L29479.1 GI:487869  
 KEYWORDS protein kinase; serine/threonine kinase.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCES 1 (bases 1 to 3447)  
 Fode, C., Motro, B., Yousefi, S., Hefernan, M. and Drenth, J. W. Sak, a murine protein-kinase/threonine kinase that is related to the Drosophila polo kinase and involved in cell proliferation Proc. Natl. Acad. Sci. U.S.A. 91 (14), 6388-6392 (1994)  
 JOURNAL MEDLINE 94294387  
 PUBMED 8022793  
 COMMENT Original source text: Mus musculus (strain DBA/2) Lymphoid tumor cDNA to mRNA.  
 FEATURES  
 source location/Qualifiers  
 1..3447 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="DBA/2"  
 /db\_xref="taxon:10090"  
 /cell\_line="D3JW25"  
 /cell\_type="Transformed cell"  
 /tissue\_type="Lymphoid tumor"  
 1..3447 /gene="sak-a"  
 1..205 /gene="sak-a"  
 206..2983 /gene="sak-a"  
 /note="sak-a and sak-b are alternate splices with divergent COOH ends beginning at nucleotide 1454; putative"  
 /product="serine/threonine kinase"  
 /protein\_id="AAC37648.1"  
 /db\_xref="GI:487870"  
 /translation="MACTGRIEDPKKGNILGKSPAGYRAISHTGLVATIMIDKKAMTKAGVORVONVEKIKOLQKPSVLELNTYFEDNNVYLVLEKCHGEMKRYLK NRMKPFSEARHFMHQIITGMVLHSHGILHRDLTNSILLTRMNLIKADFGATQ LNMPEKTYLTCGTPNYISPIATRSAGLSIDISLQCMSTYLLIGRPDPDTVKV TLNKVVLADYEMPAFLSREADLIHQLLRRNPADRLSLSSVLDPHFMGRNPSPKSKDV

GTGDSMDSGHATLSTTTTASGCTSSGSLDRLRLVGPPLNFKITVEQKNNSDPS  
SGDGSNCTOMGPNBOPANSRGRGRTYEAEERPHRYLRRAHSSDRASPQSRKAT  
YVRECHSVEMLSKPRSLDENDDSHNHCLGKTPFPADQTPQKMEVQOMGNOMY  
AHGERTNEHTVSPNRDPODYDPLDTLNAMTDTRASAKNATVAVKOLSAMKY  
MSAHNHKEVMPQEPGLHPSSEQSKNRSESTLQSKPLRSTSLPLAHLKPTROK  
TKAAVSIIDSEVCEVLEBECASBEVYEVLOISDGMPLTVVYPNDRGPPLDRP  
PLEPDIISRYSFENLPEKTYWRKQYASRYLOLRSKPTKITTYFTACCLIMENSGA  
DEEVFYDAKIHKTENLIIHIEKQISYNLKNRENTSLKEEVYMDHNEGRIC  
LSLEYSIEERKSSRSFPPIIVGRKPNSTSPALAPAPVDPSCQBDASRLS  
VNSAAPPOSRLSPSTTVBEGLGHTATATGVSLSLPKSNOLKSFVKNVQWATO  
LNSGATVWPNDGSOVLVOAGVSSIYSPDQITRYGENEKLPEYIKOKIQCISLIL  
LHFSNTEPNFQ  
2984..3447  
/gene="sak-a"

## 3'UTR

## ORIGIN

Query Match 64.5%; Score 1879; DB 10; Length 3447;  
Best Local Similarity 80.9%; Pred. No. 0;  
Matches 2362; Conservative 0; Mismatches 410; Indels 147; Gaps 9;

QY 1 ATGGCGACCTGCATCGGGAGAGATCGAGATTTTAAAGTTGGAAATCGCTGTTAAA 60  
DB 206 ATGGCGGGCGTGCATCGGGAGAGATCGAGATTTTAAAGTTGGAAATCGCTGTTAAA 265  
QY 61 GGATCATTTGCTGGTGTCTTACAGAGCTGATGTCATTCACACTGGTTTGAAGTTGCAATC 120  
DB 266 GGATCATTTGCTGGTGTCTTACAGAGCTGATGTCATTCACACTGGTTTGAAGTTGCAATC 325  
QY 121 AAAATGATAGTAAAGAAAGCATGTACAAAGCAGAAATGGTACAGAAAGTCCAAATGAG 180  
DB 326 AAAATGATAGTAAAGAAAGCATGTACAAAGTGGAAATGGTACAGAAAGTCCAAATGAG 385  
QY 181 GTGAAATATCATTTGCCAATTGAAACATCTCTTCTATCTTTGAGCTTTTAACTATTTTGA 240  
DB 386 GTGAAATATCATTTGCCAATTGAAACATCTCTCTGTCTTGGAGCTCTATATATCTTTGA 445  
QY 241 GATAGCAATTATGTATCTGTATTTAGAAATGTGCCATTAATGGAATAATGAAACAGTAT 300  
DB 446 GATAGCAATTATGTCTACCTGTATTTGGAATGTGCCAATGGAATAATGAAACAGATAT 505  
QY 301 CTAAAGAAATAGAGGAAACCTCTTCTCAGAAATGAAAGCTCGACACTTCATGACACGATC 360  
DB 506 CTGAGAAACAGAAATGAAAGCTTCTCTCAGAAAGGAAAGTAGGCACTTCATGACACGAT 565  
QY 361 ATCAAGAGATGTGTATCTTCAATTCATGATATCTACACGGAGACTTCACACTTCT 420  
DB 566 ATCAAGAGATGTATATCTTCAATTCATGATATCTACACGGAGACTTCACACTTCT 625  
QY 421 AACTCTTACTGCTGTATATGAAACATCAAGATTTGATTTTGGCTGGCACTCAA 480  
DB 626 AACTCTTACTTACGGGAATATGAAACATTAATAATTTGCTGACTTTGACTGACAAAGAG 685  
QY 481 CTGAAATATGCATGAAAGACATATATATATATGAAATCTCTTCACTCACTTTGACCA 540  
DB 686 TTGAATATGCATGAAAGACATATATATATGAAATCTCTTCACTTCACTTTGACCA 745  
QY 541 GAAATTCGCACTGAAAGTGCACATGAGCTTGAATCTGAATGTGTGGTCCCTGGCGTGA 600  
DB 746 GAAATTCGCACTGAAAGTGCACATGAGCTTGAATCTGAATTTGGTCACTGGCTGTATG 805  
QY 601 TTTTATACATTTACTATGCGGAGACACACCTTGCACATGACACAGTCAAGAACATTA 660  
DB 806 TCTTATACATTTACTATGCGGAGACACACCTTGTGACATGACACAGTCAAGAACATTA 865  
QY 661 AATAAAGTATGTTGGCAGATTTAGAAATGCCATCTTTTGTCAATAGAGGCCAAGAG 720  
DB 866 AACAAGAGTATCTGGCAGATTTAGAAATGCAAGCTTTTGTGACAGAGGCCAAGAG 925  
QY 721 CTATATTCACAGTTACTCTGTAGAAATCGACAGATGCTTTAAAGTCTGTCTTCAAGATTC 780  
DB 926 CTATATTCACAGTTACTCTGTAGAAATCGACAGATGCTTTAAAGTCTGTCTTCAAGATTC 985  
QY 781 GACCATCTTTTATGTCCGAATCTTCAACAAAAGTAAAGTTAGAACTGTGGA 840

DB 986 GACCATCTTTTATGTCCGAATCTTCAACAAAAGTAAAGACTAGGACTGTAGAG 1045  
QY 841 GACTCATTTGATAGTGGCATGCGACATATTTGATCGCAATTAACGCTTTCGATACC 900  
DB 1046 GACTCATTTGATAGTGGCATGCGACATATTTGATCGCAATTAACGCTTTCGATACC 1105  
QY 901 AGTAAAGTGATAGTTATTTTGAACAAAGACATTTTGTGTGACGACCTCCCAAT 960  
DB 1106 AGTAAAGTGATAGTTATTTTGAACAAAGACATTTTGTGTGACGACCTCCCAAT 1162  
QY 961 AAAATGATGATATTTTCAAAAGATTAAGTTCAACTGATTTTCTTTCAGAGATGGA 1020  
DB 1163 AAAATGATGATATTTTCAAAAGATTAAGTTCAACTGATTTTCTTTCAGAGATGGA 1219  
QY 1021 AACAGTTTATATCTCAGTGGGAAAT-----CAGAAACAGTAAATAGTGAAGGGA 1074  
DB 1220 AACAGTTTATATCTCAGTGGGAAATCAGAAACAGTAAATAGTGAAGGAGCGGGG 1279  
QY 1075 AAGATTAATTCAGATGAGAAAGGACCAATCTGATACCTTCTGTAAGCTTATTC 1134  
DB 1280 AAGATTAATTCAGATGAGAAAGGACCAATCTGATACCTTCTGTAAGCTTATTC 1339  
QY 1135 TCTGATAGATTCGGCACTTCTAATAGTCACTCAGCAAAACATATACATGGAACGA 1194  
DB 1340 TCTGATAGATTCGGCACTTCTAATAGTCACTCAGCAAAACATATCTAGTGAACGT 1396  
QY 1195 TGTCACTCAGCAGAAATGCTTTCAGTGCAGTCCAAAGATCAGAGAGAGTGAATGAAG 1254  
DB 1397 TGTCACTCAGCAGAAATGCTTTCAGTGCAGTCCAAAGATC----- 1434  
QY 1255 AGGATCTCAGCAGAAACATGCGCAATTTTAACTTTTAAAGAAAGACATCC 1314  
DB 1435 ----- 1434  
QY 1315 AGTATTTCTGATCTTTTGAAGACCTGATTAACATCAACACTCTCCATCATCTTTGT 1374  
DB 1435 -----ACTGATGAAATCAACAGTTCATCATCATTTGT 1471  
QY 1375 CAGGAAATATCTCTTTCATTTGACAGCCGACACTCAGATGAAACGTTACAGAG 1434  
DB 1472 CTGAGAAATATCTCTTTCATTTGACAGCCGACACTCAGATGAAATGTTACAGAG 1531  
QY 1435 TGGTTGGGAATGCAATTAATATGCTCATTTAAGAAACATATGAAATGACAGATC 1494  
DB 1532 TGGTTGGGAATGCAATTAATATGCTCATTTAAGAAACATTAAGTACACACAGCTT 1591  
QY 1495 AGCCCAACCGGACTTTCAGGCGCATCCAGATTTGCAAGAGACATCAAAAAATGCC 1554  
DB 1592 AGCCCAACAGAGATTTCAAGACATATCAGATTTGC---AGGACAGTTACGAAACGCT 1648  
QY 1555 TGGACTGATTAACAAAGTCAAAAAGATCTGTATGCTTGTATATGCACTTTGTAAA 1614  
DB 1649 TGGACTGATTAACAGAGGCGAGCAAGATCTGTATCTTGTCAATATGTTCACTGTAAAG 1708  
QY 1615 CAGCAAAATACATGAAATATATGACTGCACTTCAAGTAAACCTGAGTAAATCCAA 1674  
DB 1709 CAGCTGAGGCAATGAAATATATGAGTGCACATCACTAAGCTTGAAGTATCCACAG 1768  
QY 1675 GAATGTGTTTTGGCTCAGATCTCTTCTTGTGAACAGCAAGACATGAGGATAGAGCCA 1734  
DB 1769 GA-----GCCGGGCTTACATCTTCATTTGACAAAGCAAGATTAAGATAGAGTGC 1822  
QY 1735 CCATGGGTTATCAAGATCTGATTAAGAGACTTCACTCCGTTGGTGTCTCAGAG 1794  
DB 1823 ACATGGGTTATCAAGAACCTTACTTAAGAGATTTTCACTCTCTGATGTCTCACAG 1882  
QY 1795 TTAATAACCAACACAGAAACCAAAAGCTGTGTGTGACATCTTGAATTCAGAGAG 1854  
DB 1883 TTAAGCCAAATCAACAGAAACCAAAAGCTGTGTGTGACATCTTGAATTCAGAGAG 1942  
QY 1855 GTGTGTGTGAGCTTTGAAGAGTATGATCTCAAGAAATATGTGAAGAGTTCTTCAG 1914



Db 1943 GTGTGTGTGAGCTTCTGAGAGAGTGTGCGCTGGAAGATATGTGAAAGAGTGTTCAG 2002  
 QY 1915 AATCTAGTGTGAAATATGATCACTATTATATCCAAATGAGTGTAGAGTTTCTT 1974  
 Db 2003 AATGTGAGTGTGAGGAGTATGATCACTGTATTATACCCCAAGATGGAAGAGCTTCT 2062  
 QY 1975 CTGTGTGTGATGACCACTCTGACCTACTGACATCATGAGGTACAGCTTTGACATTTA 2034  
 Db 2063 CTGTGTGTGACAGACCTCTGCTGTCTACTGACATCATGAGGTACAGCTTTGACATTTA 2122  
 QY 2035 CCAGAAAATATCTGCGGAAATATCAATATGCTTCCAGGTTTGTACAGCTTGTAGATCT 2094  
 Db 2123 CCAGAAAATATCTGCGGAAATATCAATATGCTTCCAGATTCATTCAGCTAGTAAATCT 2182  
 QY 2095 AATCTTCCAAATATCACTATTATTTACAAATATGCTAAATGCACTTTTGTATGGAATTC 2154  
 Db 2183 AAAATCTCCCAAAATCACTATTATTTACAAATATGCTAAATGCACTTTTGTATGGAATTC 2242  
 QY 2155 CCGTGTGTGATTTTGTAGGTTTGTGTATGATGAGGTGAAATATACAAACAGAAAGAT 2214  
 Db 2243 CCGTGTGTGATTTTGTAGGTTTGTGTATGATGAGGTGAAATATACAAACAGAAAGAT 2302  
 QY 2215 TTCACTTCACTGATGATGAAAGAGAGGAGATCTTACCTTTAAAGTAAAGTAAAGT 2274  
 Db 2303 TTCACTTCACTGATGATGAAAGAGAGGAGATCTTACCTTTAAAGTAAAGTAAAGT 2362  
 QY 2275 AATAGCTTGAAGAGAGATTAATAATGATATGATGAGATGCTTAAATGAGGTGATCTAT 2334  
 Db 2363 ACCAGCTGGAAGAGAGATTAATAATGATATGATGAGATGCTTAAATGAGGTGATCTAT 2422  
 QY 2335 TGTATTAGAGCTGGAATCCATTAATTTTCAAGAGAGAGAGAGAGAGTGTCTCCCTT 2394  
 Db 2423 TGTATTAGAGCTGGAATCCATTAATTTTCAAGAGAGAGAGAGAGAGTGTCTCCCTT 2482  
 QY 2395 TTCCCAATATATGATGAG 2454  
 Db 2483 TTCCCAATATATGATGAG 2542  
 QY 2455 CCTCTCTGTGTGATTTCAAAATATCCCAAGAGAGATGAGATCTTTCAACAGATGATC 2514  
 Db 2543 CCTCTCTGTGTGATTTCAAAATATCCCAAGAGAGATGAGATCTTTCAACAGATGATC 2599  
 QY 2515 ATGATAGTGTGCTTCTTCCAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2574  
 Db 2600 GTGAATAGTGTGCTTCTTCCAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2659  
 QY 2575 GAAAGACTTGTCTTCACTACAGCTTCTGGAACAGAGATCTTCTTAATAGTCTTAA 2634  
 Db 2660 GAAAGACTTGTCTTCACTACAGCTTCTGGAACAGAGATCTTCTTAATAGTCTTAA 2708  
 QY 2635 GATTGTCTTCTTCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2694  
 Db 2709 GATTGTCTTCTTCACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2764  
 QY 2695 GGTACAGAGTAACTAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2754  
 Db 2765 GGTACAGAGTAACTAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2824  
 QY 2755 GTGACAG 2814  
 Db 2825 GTGACAG 2884  
 QY 2815 GGAAGAAATGAAATATACAG 2874  
 Db 2885 GGAAGAAATGAAATATACAG 2944  
 QY 2875 CTTTGTGATTTTCTTAATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2913  
 Db 2945 CTTTGTGATTTTCTTAATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2983

RESULT 9  
 BC026785

**LOCUS** BC026785 3473 bp mRNA linear ROD 06-OCT-2003  
**DEFINITION** Mus musculus serine/threonine kinase 18, mRNA (CDNA clone MGC:30237 IMAGE:5150173), complete cds.  
**ACCESSION** BC026785  
**VERSION** BC026785.1 GI:20072408  
**KEYWORDS** MGC.  
**SOURCE** Mus musculus (house mouse)  
**ORGANISM** Mus musculus  
**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**AUTHORS** Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, C.M., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farnet, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Umed, T.B., Toshimiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalobon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pahey, J., Helton, E., Kettleman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Kiryushinski, M.I., Skalska, U., Smalhe, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Matra, M.A., Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
**TITLE** human and mouse cDNA sequences  
**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
**MEDLINE** 22388257  
**PUBMED** 12477932  
**REFERENCE** 2 (bases 1 to 3473)  
**AUTHORS** Strausberg, R.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (04-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
**REMARK** NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
**COMMENT** Contact: MGC help desk  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loubege, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.  
**FEATURES**  
 source  
 1..3473  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="MGC:30237 IMAGE:5150173"  
 /tissue\_type="mammary tumor. MAP-2F alpha model. 7 months old, gross tissue."  
 /clone\_id="NCI\_CGAP\_Mam5"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMD ac: <http://image.llnl.gov>  
 Series: IRAX Plate: 41 Row: P Column: 11  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27501453.  
 Location/Qualifiers  
 1..3473  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="MGC:30237 IMAGE:5150173"  
 /tissue\_type="mammary tumor. MAP-2F alpha model. 7 months old, gross tissue."  
 /clone\_id="NCI\_CGAP\_Mam5"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

```
gene
1..3473
/ gene="Stk18"
/ note="synonym: 1700028H20"
/ db_xref="LocusID:20873"
/ db_xref="MGI:101783"
214..2991
/ codon_start=1
/ product="serine/threonine kinase 18, isoform sak-a"
/ protein_id="AAH26785.1"
/ db_xref="GI:20072409"
/ db_xref="LocusID:20873"
/ translation="MAACIGERIEDPKVGNLKGSPAGVYRAESIHTELEVALIKMID
KKMYKAGVQVQNEVKIHQDKPSVLELYNYPEDNNVYVYLMECHNGENRILK
NMKPPSERARHPHQIITGMYLHSHGIDHDLTSLNLTFRNNKILADPGALTO
LNMHPKHVTLCTPNVYIPEIATRSAGHLESDIWSLGCMEFTLLIGRPPIPDVKN
TLNKLVLADYEMPAFLSRAQDLIHLRLNRPADRLSLISVDLPHMSNRPSPKSDV
GTVEDMSDGHATLSTTITASSGSTSLSGSLDBRLLVCPLEKTLIVFQKNKSDPS
SGGNSNFTQMKPPEOBANSRGRVYEDAEKPSRYLRRAHSDRASPSQSRKAT
YSVERCHSVEMLSKPRSLDENQSNHHCLEKTPPPADVOTPOEMVQWQENLQMN
AHLEGTNEHTVSPNFDYDPLQDILNANMTDTRASKADTASAVNAVHAKOLAMKY
MSAHHRKPEVMPQEPDLHPHSEQSKNSMESTLGYOKPLRLSTISPLIARLKPTRK
TKKAVVSLDSEVEVLEKRESEGVYEVLOISDGMITVYVYNDGRGPLADRP
PLPTNNISRYSPNLPKRYKRYQVARSPIQLVRSKTPKITTFTRIAKCLIMENPGA
DEWVRYDCAKIKHTENLHILKGTISLNLKNEVSLKKEVAKVMDHANGRLIC
LSLESVISSEKRSRGSFPLIVGRKPNSTSPKALSLKPSADLKSYPVKNVWATO
VNSAAPVTPSGPLSPSTVVEG/GHTAATGTGVSLSLPSKADLKSYPVKNVWATO
LTSGAVVQFNDGSQLVQAGVSSISYTSPPDQTRRYGENEKLPEYIKQKCLSSIL
LMESNTPNPO"
247..1008
/ note="S. Tkc, Region: Serine/Threonine protein kinases,
catalytic domain"
/ db_xref="CDD:smart0220"

misc_feature
Query Match 64.3%; Score 1874.2; DB 10; Length 3473;
Best Local Similarity 80.8%; Pred. No. 0;
Matches 2359; Conservative 0; Mismatches 413; Indels 147; Gaps 9;

ORIGIN
1 ATGGCGACTGCATCGGGAGAGAAATCGAGATTTTAAAGTTGAAATCTGCTGTGTA
214 ATGCGCGCGTGCATCGGGAGAGATCGAGACTTTAAAGTTGAAATCTACTCGTAA
61 GGATCATTTGCTGTGCTACAGAGCTGATTCATTCACATCGTTTGAAGTTGCAATC
274 GATCATTTGCTGTGCTACAGAGCTGATTCATTCACATCGTTTGAAGTTGCAATC
121 AAAATGATGATGAAGAACCATGTACAAGCAGAGATGTACAGAGATCCAAATGAG
334 AAAATGATGATGAAGAACCATGTACAAGCTGGAATGTACAGAGATCCAAATGAG
181 GTGAAGAAATCATTCGCAATTTGAACAATCTTATCTTGAGCTTTATTAATTTGAA
394 GTGAAGAAATCATTCGCAATTTGAACAATCTTATCTTGAGCTTTATTAATTTGAA
241 GATAGCAATTTATGTGTATCTGTATTAAGAAATGTCATTAATGAGGAATGAACGAT
454 GATTAACATTTATGTGTATCTGTATTAAGAAATGTCATTAATGAGGAATGAACGAT
301 CTAAAGATGAGTGAACCTCTTCTAGAAAATGAAGCTCGACATTCATGACCAAGATC
514 CTGAAGAACAGAAATGAGCTTTCTAGAAAAGGAAAGCTAGGCACTTATGACCAAGAT
361 ATCAGAGGATGTGTATCTTATCTGTATCTGTATTAAGCAACCGGACCTTCACTTTCT
633 ATCAGAGGATGTGTATCTTATCTGTATCTGTATTAAGCAACCGGACCTTCACTTTCT
421 AACCTCTACTGACTCTGTATTAAGCAACATGAAGTTGTGATTTTGGGCTGGCAACTCA
480 AACATCTTACTTAAAGCGGAAATATGAACATTAATAATTTGTGACTTTGACCTGCAACG
634 AACATCTTACTTAAAGCGGAAATATGAACATTAATAATTTGTGACTTTGACCTGCAACG
481 CTGAAGATGCAATGAAGACATTAACATTAATGTGAACTCTTAATCAATTTCAACA
540 TTGAATATGCAATGAAGACATTAACATCTGTGTGGACTCTTAATTAATTTATCAACA
694 TTGAATATGCAATGAAGACATTAACATCTGTGTGGACTCTTAATTAATTTATCAACA
753
```

```
541 GAAATTCGCACTGGAAGTGCATGAGCCCTTGAATCTGATGTTGGTCCCTGGGCTGTATG
600
754 GAAATTCGCACTGGAAGTGCATGAGCCCTTGAATCTGATGTTGGTCCCTGGGCTGTATG
813
601 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
660
814 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
873
661 AATTAAGATGATTTGGCAGATTTATGAATGCAATCTTTTGTCAATAGAGGCCAAGGAC
720
874 AATTAAGATGATTTGGCAGATTTATGAATGCAATCTTTTGTCAATAGAGGCCAAGGAC
933
721 CTATATTCACAGTTACTCTTGTGAATTCAGACAGATGTTTAACTGTCTTCAAGATTTG
780
934 CTATATTCACAGTTACTCTTGTGAATTCAGACAGATGTTTAACTGTCTTCAAGATTTG
993
781 GACCATCTTTTATGTCGGAATCTTCAACAAAAGTAAAGATTTAGAGGACCTGTGAA
840
994 GACCATCTTTTATGTCGGAATCTTCAACAAAAGTAAAGATTTAGAGGACCTGTGAA
1053
841 GACTCAATGATGATGAGGACATGCAATTTCTACTGCAATTAAGCTTCTTCAAGTACC
900
1054 GACTCAATGATGATGAGGACATGCAATTTCTACTGCAATTAAGCTTCTTCAAGTACC
1113
901 AGTAAAGTGTAGTTATTTGACAAAAGAAAGACTTTGATGTCAGCCATCCCAAT
960
1114 AGTAAAGTGTAGTTATTTGACAAAAGAAAGACTTTGATGTCAGCCATCCCAAT
1170
961 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
1020
1171 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
1227
1021 AACGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1074
1228 AGTAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1287
1075 AGAGTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1134
1288 AGAGTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1347
1135 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1194
1348 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1404
1195 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1254
1405 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1442
1255 AGGTACTCACCACAGACAAATGACCAATCTTTTAAAGAAAGAACATCC
1314
1443 -----
1442
1443 -----
1442
1315 AGTAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1374
1443 -----
1442
1375 CAAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1434
1480 CTAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1539
1435 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1494
1540 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1599
1495 AGCCCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1554
1600 AGCCCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1656
1555 TGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1614
1657 TGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
1716
```

1615 CAGCAAAATACATGAAATATATGACTGCACTTCAAGTAACCTGAGATTAATCAACA 1674  
 1717 CAGCTGAGTCCATGAAATACATGATGCAATCACTAAGCTGAGATGCAAG 1776  
 1675 GAATGTTTTTGGCTCAGATCTCTTTCTGAAAGAGCAAGATGAGGGTATGAGCCA 1734  
 1777 GA-----GCCGACCTTATCATCTCTGAAAGCAAGCAAGATGAGTATGAGTGG 1830  
 1735 CCATGGGGTTATCAGAAATGCTATTAAGAGATTAATCCCTGGTGGTCTCAG 1794  
 1831 AACCTGGGTTTACAGAAACCTTAAAGATTAATCACTCTCTGATGCTCACA 1890  
 1795 TTAATAACCATCAGACAGAAACCAAAAGGCTGTGTGAGCATCTTGATTCAGAGAG 1854  
 1891 TTAAGCCATCAGACAGAAACCAAAAGGCTGTGTGAGCATCTTGATTCAGAGAG 1950  
 1855 GTGTGTGTGAGCTGTGAAAGAGTATGCAATCAAGAAATGTAAGAAAGTTCTTCAG 1914  
 1951 GTGTGTGTGAGCTGTGAAAGAGTATGCAATCAAGAAATGTAAGAAAGTTCTTCAG 2010  
 1915 ATATCTAGTATGAAATACGATCACTATTTATTAATCCAAATGTTGATGAGTTTCT 1974  
 2011 ATATCTAGTATGAAATACGATCACTATTTATTAATCCAAATGTTGATGAGTTTCT 2070  
 1975 CTTGCTGATGAGCAACCTCCTCACTGACCAATCATGATGATGATGATGATGATGAT 2034  
 2071 CTTGCTGAGACACTCTCTCTGCTTACTTAAACATCATGATGATGATGATGATGAT 2130  
 2035 CCAGAAATACTGCGGAAATATCAATATGCTTCCAGGTTGTACAGCTTGTATGATCT 2094  
 2131 CCAGAAATACTGCGGAAATATCAATATGCTTCCAGGTTGTACAGCTTGTATGATCT 2190  
 2095 AAATCTCCCAATATCTTATTTTACAAATATGCTTAAATGCTTTGATGAGATTTCT 2154  
 2191 AAAAATCCCAATATCTTATTTTACAAATATGCTTAAATGCTTTGATGAGATTTCT 2250  
 2155 CCGGCTGATGATTTTATGAGTTGTTTATGATGAGGTTTAAATATACAAAGAT 2214  
 2251 CCGGCTGATGATTTTATGAGTTGTTTATGATGAGGTTTAAATATACAAAGAT 2310  
 2215 TTCAATCAGGTATGAAAGAGAGGAGTCTTACACTTAAAGTGAAGTGAAGTT 2274  
 2311 TTAATTCATATATGAAAGAGAGGAGTCTTATTAATTAATTAATTAATTAATTAAT 2370  
 2275 AATAGCTGAAAGAGAGATTAATATGATATGATGATGATGATGATGATGATGAT 2334  
 2371 ACCAGCTGAAAGAGAGATTAATATGATATGATGATGATGATGATGATGATGAT 2430  
 2335 TGTTTAGCATGGAATCCATATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2394  
 2431 TGTTTAGCATGGAATCCATATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2490  
 2395 TTCCCAATATATCATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2454  
 2491 TTCCCAATATATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2550  
 2455 CCTCTCTCTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2514  
 2551 CCTCTCTCTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2607  
 2515 ATGCAATAGTGTCTCTCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2574  
 2608 GTGAATAGTGTCTCTCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2667  
 2575 GAAGAGCTGTGTCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2634  
 2668 GAAGAGCTGTGTCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2716  
 2635 GATGTCTCTCTTCAACAG 2694  
 2717 ---GTTCTCTTCAACAG 2772  
 2695 GCTACACAGTAACTAGT 2754

2773 GCTACACACTAATGAGCGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2832  
 2755 GTGACAGCAGAGAGTGTCTTATGATGATGATGATGATGATGATGATGATGATGAT 2814  
 2833 GTTCAGAGCAGAGAGTGTCTTATGATGATGATGATGATGATGATGATGATGATGAT 2892  
 2815 GGAGAAATGAAATATTAACAGATCATCAACAGAAATTAACAGATGATGATGATGAT 2874  
 2893 GGAGAAATGAAATATTAACAGATCATCAACAGAAATTAACAGATGATGATGATGAT 2952  
 2875 CTTTGATGTTTCTTAATCGAATCTCTTAATTTTATGTA 2913  
 2953 CTTTGATGTTTCTTAATCGAATCTCTTAATTTTATGTA 2991

RESULT 10  
 BC057940  
 LOCUS  
 DEFINITION  
 Mus musculus serine/threonine kinase 18, mRNA (CDNA clone MGC:65368  
 IMAGR:5323295), complete cde.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 3377)  
 Strausberg, R.L., Feldgould, B.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, P.S., Wagner, L., Shemmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Mang, J., Hsieh, F.,  
 Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Uebli, T.B., Toshitoki, S.,  
 Carinci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, B.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,  
 Schermer, A., Schein, J.E., Jones, S.J., and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 22388257  
 JOURNAL  
 MEDLINE  
 PUBMED  
 12477932  
 2 (bases 1 to 3377)  
 Strausberg, R.  
 Direct Submission  
 Submitted (08-SEP-2003) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Jeffrey Green M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nsl.nih.gov](mailto:nisc_mgc@nsl.nih.gov)  
 Akhtar, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,  
Machiro, Q. L., Masello, C., Maskell, B., Mastrian, S. D., McCloskey, J. C.,  
McDowell, J., Pearson, R., Skantirlop, S., Thomas, P. J., Touchman, J. W.,  
Tsurgren, C., Vogt, J. L., Walker, M. A., Weeber, K. D., Wiggins, L.,  
Young, A., Zhang, J. H. and Green, E. D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
Series: IRAC Plate: 123 Row: 1 Column: 15  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 27501453.

## FEATURES

## SOURCE

1..3377

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="MGC:53368 IMAGE:5333295"

/issue\_type="Mammary tumor, C3(1)-Tag model, Infiltrating

ductal carcinoma, 5 month old virgin mouse."

/clone\_1db="NCI CGAP\_Mam6"

/lab\_host="DH10B"

/note="Vector: PCMV-SPORT6"

1..3377

/gene="Skkl8"

/note="Synonym: 1700028H20"

/db\_xref="LocusID:20873"

/db\_xref="MGI:101783"

201..2897

/product="Skkl8 protein"

/protein\_id="AAH57940.1"

/db\_xref="GI:3478391"

/db\_xref="LocusID:20873"

/transcript="MAKIGER1EDPKVGNLKGSPAGVRAESIHTEJAVAKMID

KAMVAVGVOROVMEKIKHQLKPSVLELYNEDNNVYVLIEMGNGMNVYK

NMKPFSEERAEHFMHIOITGMILYHSHGLHRDLITLITRMNITKAPGATQ

LNNHKEHYTLGGTTPHUIPEIATRSAGHLESDISGCMPTTILIGRPFDIVKN

TINKVYLADYEMAPLISRAOPLIHOLLRNPADRLSSVLDHPMSRNPSPKSDV

GTVESBMDSGHATLSTTTTASGSLSSGLDLDRLLVQPLPNKLTIVQKNKSDFS

SEDGNSFTQWGNPEQANSRGRVLEDAERPHSRYLRAHSDRASPSNOSAKT

YSEVERCHSEVMSKPRSLDENQSHNHCLGTPPEPADOTPOKEMVOQWGNLQNM

AHIGTNEHTVSPNRFQDYLDTLLENAWTDRASKNADTSANVAIVQNSMKY

MSAHHKPEVMEPOEDLHPHSEKSNRSMESTLGYOKTSLRIS7SPILAHKLPIROK

RILQVRSCTPKITITTRAKTILMENSADFEVFDAGAIHTEMLIHIETGIGI

SINLKNENVTLSKEBKVYMDHANEGRILCLSESVLSBEKRSSTSPVPIIVGRK

PENTSSPKLASLPVDPSCCKGBOASASRLISVNSAAPTOSBEKRSSTSPVPIIVGRK

TATGTGVSLSLPSKSQLKSVKVNQWATOLSGAVWVQFNDGQLVVOAGVSISV

TSPDQTTTGYGENEKLPEYIKOKLQCLSSILLMSPNTPNFQ"

234..995

/note="S\_TKC; Region: Serine/Threonine protein kinases,

catalytic domain"

/db\_xref="CD:smart00220"

ORIGIN

## misc\_feature

/note="S\_TKC; Region: Serine/Threonine protein kinases,  
catalytic domain"

/db\_xref="CD:smart00220"

Query Match 59.2%; Score 1724.6; DB 10; Length 3377;  
Best Local Similarity 78.5%; Pred. No. 0;

Matches 2292; Conservative 0; Mismatches 399; Indels 228; Gaps 10;

OY 1 ATGGGACCTGCATGGGAGAGATGAGATTTTAAAGTTGGAATCTGCTGTAA 60  
DB 201 ATGGGGGGTGCATGGGAGAGATGAGACTTTAAGTTGGAATCTGCTGTAA 260  
OY 61 GGATCATTTGCTGGGTCTACAGAGCTGAGTCCATTCACCTGTTTGAAGTTGCAATC 120  
DB 261 GGATCATTTGCTGGGTCTACAGAGCTGAGTCCATTCACCTGTTTGAAGTTGCAATC 320  
OY 121 AAAATGATGATGAAGAAAGCCATGTACAAAGCAGAAATGTACAGAGAGTCCAAATAG 180  
DB 321 AAAATGATGATGAAGAAAGCCATGTACAAAGCAGAAATGTACAGAGAGTCCAAATAG 380  
OY 181 GTGAATATACATTCGCAATGAACATCTTCTTATCTTGAAGCTTTATTAATATTTGAA 240

DB 381 GTGAATATACATTCGCAATGAACATCTTCTTATCTTGAAGCTTTATTAATATTTGAA 440  
OY 241 GATAGCAATTAATGTATCTGTGATTAAGAAATGTCATTAAGGAAATGAACGATAT 300  
DB 441 GATAGCAATTAATGTATCTGTGATTAAGAAATGTCATTAAGGAAATGAACGATAT 500  
OY 301 CTAAAGATGATGATGAAGAAAGCCCTTCTCAGAAATGAAGTTCGACACTTCATGACAGATC 360  
DB 501 CTGAAGAAACAGATGAAGCCCTTCTCAGAAAGGAAAGTGAAGCACTTCATGACAGAT 560  
OY 361 ATCAGAGGATGTGATCTTCATCTCATGTGATATCAACCGGAGACTCACTTCT 420  
DB 561 ATCAGAGGATGTGATCTTCATCTCATGTGATATCAACCGGAGACTCACTTCT 620  
OY 421 AACTCTTACTGCTGTAATATGAAATCAATCAAGATTCGATTTTGGCTGGCACTCA 480  
DB 621 AACATCTTACTTACCGGAAATATGAACATTAATGCTGACTTTGACATGACAGAG 680  
OY 481 CTGAAGAAATGCCATGAAGAAAGCATATACATTAATGTGAACTCTTAATCAATTCACA 540  
DB 681 TTGAATATGCAATGAAGAAAGCATATACATCTGTGGACTCTTAATATATTCACA 740  
OY 541 GAAATTCGCACTCGAAGTGCATGAGCCCTTGAATCTGATGTTTGGTCCGAGCTGATG 600  
DB 741 GAAATTCGCACTCGAAGTGCATGAGCTTGAATCTGATGTTTGGTCCGAGCTGATG 800  
OY 601 TTTTATACATTAATGATGAGGAGCAACCTTGGACATGACAGATCAAGAACATTA 660  
DB 801 TTTTATACATTAATGATGAGGAGCAACCTTGGACATGACAGATCAAGAACATTA 860  
OY 661 AATAAGATGATGATGAGGAGATTAAGAAATGCAATCTTTTGTCAATAGAGCCAGAG 720  
DB 861 AATAAGATGATGATGAGGAGATTAAGAAATGCAATCTTTTGTCAATAGAGCCAGAG 920  
OY 721 CTTAATGCAAGATTAATGATGAGGAGATGCAAGATGATGATGATGATGATGATG 780  
DB 921 CTTAATGCAAGATTAATGATGAGGAGATGCAAGATGATGATGATGATGATGATG 980  
OY 781 GACATCTCTTAAATGTCGCAAAATCTTCAACAAAGATTAAGATTAAGATTAAGAT 840  
DB 981 GACATCTCTTAAATGTCGCAAAATCTTCAACAAAGATTAAGATTAAGATTAAGAT 1040  
OY 841 GACTCAATGATGATGAGGAGATGCAATCTTCACTGCAATTAAGATTAAGATTAAG 900  
DB 1041 GACTCAATGATGATGAGGAGATGCAATCTTCACTGCAATTAAGATTAAGATTAAG 1100  
OY 901 AGTATAGTGTGATGATTAATGCAAAAGAGCTTTGATGATGATGATGATGATG 960  
DB 1101 AGTATAGTGTGATGATTAATGCAAAAGAGCTTTGATGATGATGATGATGATG 1157  
OY 961 AAAATGATGATGATGATGATTAAGAAATGCAATCTTCTTCAAGAGATGATG 1020  
DB 1158 AAAATGATGATGATGATTAAGAAATGCAATCTTCTTCAAGAGATGATGATG 1214  
OY 1021 AACAATTTTAAATCTGATGAGGAGAAAT-----CAAGAAACAGTAAATGATGAGGAG 1074  
DB 1215 AACAATTTTAAATCTGATGAGGAGAAATCAAGAAACAGTAAATGATGAGGAGG 1274  
OY 1075 AGAGTAATCAAGATGAGGAGAAAGCCATCTTCATGATGATGATGATGATGATG 1134  
DB 1275 AGAGTAATCAAGATGAGGAGAAAGCCATCTTCATGATGATGATGATGATGATG 1334  
OY 1135 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1194  
DB 1335 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1391  
OY 1195 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1254  
DB 1392 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1429  
OY 1255 AGGATCTACCAAGAGCAATGCAATTTTAACTTCTTAAAGAAAGATCTC 1314

Db 1430 ----- 1429  
 Qy 1315 AGTAGTCTGATCTTTGAAAAGCCTGATTAACAATGACACTCTCCATCATCTTTGT 1374  
 Db 1430 -----ACTGATGAAAATCAACACAGTTCCAAATCATCTTGT 1466  
 Qy 1375 CCAGGAAAACTCTTTTCATTTGACAGCCGACACTCAGACTGAAACCGTACACAG 1434  
 Db 1467 CTAGGAAAACTCTTTTCATTTGACAGCCGACACTCAGACTGAAATGTTACACAG 1526  
 Qy 1435 TGGTTGGGAATCTGCAATTAATGCTCATTTAAGAAAACTACTGAAATATGACAGCATC 1494  
 Db 1527 TGGTTGGGAATCTGCAATTAATGCTCATTTAAGGAAATATGACAGCCAGCGTT 1586  
 Qy 1495 AGCCCAACCGGAGCTTCCAGGCGCATCCAGATTGCGAAGAGACATCAAAAAATGCG 1554  
 Db 1587 AGCCCAACAGAGATTTCCAGAGCATATCCAGATTTGCG---AGAGACAGTTTACGAAACGCT 1643  
 Qy 1555 TGGAGCTGATACAAAGTCAAAAAGAACTGATGCTTGTATTAATGCAATTCGTAA 1614  
 Db 1644 TGGAGCTGACAGAGCCGACAGAAATGCTGATCTTGTCCAAATGTTTCACTGTAAAG 1703  
 Qy 1615 CAGCAAAATACCATGAAATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAA 1674  
 Db 1704 CAGCTGAGTGCATGAAATATACATGAGTGCACATCATTAAGCTGAGGTCAATGCCAG 1763  
 Qy 1675 GAATGTTTGGTCTCAGATCTCTTTCTTGAACAGACAGACTAGGGGTATGAGCCA 1734  
 Db 1764 GA-----GCCGAGCTTACATCTCTCATTCGAAACAAAGCAAGATGAAATGAGGTG 1817  
 Qy 1735 CGATGGGGTTTATCAGAAATGATATTAAGAGCATTAATCTCGTTGGTGTCCACAG 1794  
 Db 1818 ACATGGGGTTTATCAGAAATCTTAAAGATTAATCACTCTCTGATGTTGTCACAGA 1877  
 Qy 1795 TTAATAACCAATGACAGAGAAAAACAAAAAGCTGTGTGTGAGCATCTTGAATTCAGAGAG 1854  
 Db 1878 TTAAGGCAATGACAGAGAAAAACAAAAAGGCTGTGTGTGAGCATCTTGAATTCAGAGAG- 1936  
 Qy 1855 GTGTGTGTGAGCTGTGTAAGAGATGCAATCTCAAGAAATATGTGAAGAAGTTCTTCAG 1914  
 Db 1937 ----- 1936  
 Qy 1915 AATATCTAGATGAGAAATAGCATCTATTTATTCAAATGAGTGTAGAGTTTCT 1974  
 Db 1937 -----GATCAGCTGTTTATTTACCCGAAACAGTGAAGAGCTTTCT 1976  
 Qy 1975 CTGTGATATGACCAACCTCACCCTACTGCAACATCATGATGATGAGCTTTGCAATTTA 2034  
 Db 1977 CTGTGATGACAGCTCCTTGCCTTCTTAACAACATCATGATGATGAGCTTTGCAATTTA 2036  
 Qy 2035 CCAGAAAAATCTGCGGAAAAATATCAATATGCTTCCAGGTTGTACAGCTTGTAGATCT 2094  
 Db 2037 CCAGAAAAATCTGCGGAAAAATATCAATATGCTTCCAGGTTGTACAGCTTGTAGATCT 2096  
 Qy 2095 AATCTCCCAAAATCACTATTTTACAAGATATGCTAAATGCACTTTTGTAGAGATTTCT 2154  
 Db 2097 AAAAACTCCCAAAATCACTATTTTACAAGATATGCTAAATGCTTTTGTAGAGATTTCT 2156  
 Qy 2155 CCGTGTGCTGATTTTGAAGTTTGGTTTATGATGAGGGTAAATAACAAACAGAAAT 2214  
 Db 2157 CCGTGTGCTGATTTTGAAGTTTGGTTTATGATGAGGGTAAATAACAAATGAAAT 2216  
 Qy 2215 TTCAATTCAGGTGATTTGAAAAGACAGGAAAGTCTTACCTTTAAAAAGTGAAGTT 2274  
 Db 2217 TTAAATTCACATATTTAGAAAACAGGAAATCTTAAATTTAAAAATGAAAATGAAGTT 2276  
 Qy 2275 AATAGCTTGAAGAGAGATTAATAATGATATGATGAGCATCTGATAGGGGTATGAT 2334  
 Db 2277 ACCAGCTTGAAGAGAGATTAATAATGATATGATGAGCATCTGATAGGGGTATGAT 2336  
 Qy 2335 TGTATGACATGATATCAATATTTTGAAGAGAAAAAGAAAACTAGAGATGCTCCCTTT 2394  
 Db 2337 TGTATGACATGATATCAATATCTCTGAGAGAGAAAAAGAAAAAGGGGTCTTCAATTC 2396

Qy 2395 TTCCCAATATCATAGAGAAAACTGGTAGTACTAGTTCACTTAAGGCTTATCACCT 2454  
 Db 2397 TTCCCTATATATGATGAGAGAAAACTGGTATATCTAGTTCACTTAAGGCTTATCACCT 2456  
 Qy 2455 CCTCTCTGTGGAATCAAAATTAACCAACGAGAGATAGAGATCTTTCAACGAATGCTC 2514  
 Db 2457 CCTC---CTGTGAGCCCAAGCTGTATGAAGAGAGAGAGGCGCTCAGCAAGAGACTGAGC 2513  
 Qy 2515 ATGCATATGCTGCTCTTTCTTCAACAGAGCAACCAATCCTTAATCCCTATGTTCAAT 2574  
 Db 2514 GTGAATAGTCCGCTTTCCCAACAGATCCCAAGACTCAGTCCCTTCACTGTGACAGTT 2573  
 Qy 2575 GAAGACTGTGCTTACAACTACAGCTTCTGAGACAGACATCTCTTATATAGTATAAA 2634  
 Db 2574 GAAGACTGTGCTTACAACTACAGCTTCTGAGACAGACATCTCTTATATAGTATAAA 2634  
 Qy 2635 GATGTCTCTTCTTAAATCAAGCAACATCTTTGAAATCTGTTTGTGAAAAATGTGTTG 2694  
 Db 2632 ----GCTTCTTAAATCTGCAACAGCTTTGAAATCTGTTTGTGAAAAATGTGTTG 2678  
 Qy 2695 GCTACAGCTTAACTATGAGAGAGCTGTGTGCTTCACTTAAATGATGCTCCAGTTGTT 2754  
 Db 2679 GCTACAGCTTAACTATGAGAGAGCTGTGTGCTTCACTTAAATGATGCTCCAGTTGTT 2738  
 Qy 2755 GTGACGAGAGAGTCTTCTTATCACTTATTAACCTCAACCAATGTCTCAACACTAGAT 2814  
 Db 2739 GTGACGAGAGAGTCTTCTTATCACTTATTAACCTCAACCAATGTCTCTTCACT 2798  
 Qy 2815 GGAGAAATGAAATAATTAACAGACTTACATCAACAGAAATTAACAGTCTGTCTTCATC 2874  
 Db 2799 GGAGAAATGAAATAATTAACAGACTTACATCAACAGAAATTAACAGTCTGTCTTCATC 2858  
 Qy 2875 CTTTGTATGTTTCTTATCTGAGCTGCTTATTTTCACTGA 2913  
 Db 2859 CTTGATGTTTCTTATCTGAGCTGCTTATTTTCACTGA 2897

RESULT 11  
 BC060363  
 LOCUS 3554 bp mRNA linear VRT 21-OCT-2003  
 DEFINITION Xenopus laevis cDNA clone MGC:68791 IMAGE:4202704, complete cds.  
 ACCESSION BC060363  
 VERSION BC060363.1 GI:37805196  
 KEYWORDS MGC.  
 SOURCE Xenopus laevis (African clawed frog)  
 ORGANISM Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Xenopus.  
 1 (bases 1 to 3554)  
 Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.  
 and Richardson, P.  
 Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 Initiative  
 JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)  
 MEDLINE 22341132  
 PUBMED 12454917  
 REFERENCE 2 (bases 1 to 3554)  
 Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hale, F.,  
 Diachenko, L., Marusina, K., Ramet, A.A., Rubin, G.M., Hong, L.,  
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Uebelin, T.B., Toshilyuk, S.,  
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Millan, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultky, S.W.,  
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,



Db 1087 AGTAAATGCGCCGTTTTCAGAGAAAAGAA---TCCTGTGCTGGCCCTTCGCTCCCAAC 1143  
 Qy 961 AAAATACGTATATTCCTCAAGATTAAGTTCACTGATTTTCTTCCAGAGATGCA 1020  
 Db 1144 AAAGTCATATCTTC---AATTTAAGATTAACATCTTAAGTAAGATCAAAATGATGGA 1200  
 Qy 1021 AACAGTTTTTATCTAGTGGGGAATCAAGAAACAGTAATAGTGAAGGGGAAGTA 1080  
 Db 1201 AGCTTCTATATACAGAAAGGAAATTAATGA---CTTTCTGAGGGCAATGAAAGAAA 1257  
 Qy 1081 ATTCAAGATGACAGAAAGGACCAATTCATCTGATCTTGTAGAGCTTAATTCCTGAT 1140  
 Db 1258 CCGTGTGCTGTGAGGACAGACCAATTCGCGTTTACTGGAAAGGGCTCATCTCATGAC 1317  
 Qy 1141 AGATCTGCACTTCTAATAGTCAAGCAAAAACATATA---CAATGAAAGCATGT 1197  
 Db 1318 AGGTCAAGAACTCTCAAGTCAAACTTATGTCAACAAAGAGCTAATTCAGAGCGGTGC 1377  
 Qy 1198 CACTCAGAGAAATGCTTCAAGTGTCCAAAGATCAGAGAGGTGAAAATGAAAGAGAG 1257  
 Db 1378 CATTCAGTGAAGTCTTGTCTTAACCCATCTCATTTAAGGGCTAACCGCACT----- 1428  
 Qy 1258 TACTCAACCAAGACAAATGCAATCTTTTAACTTCTTAAAGAAAAGACATCAGT 1317  
 Db 1429 TCTTCACTCCCAACAGCTATGAGACATACACAGATGTTCAAGATGAACGTTCTTTG 1488  
 Qy 1318 AGTTCGATCTTTTGAAGAAGCTGATTAACATCAAGCACTCTCCATATCATCTTGTCCA 1377  
 Db 1489 GAGAGCACATCTCTCC---TCCAGTGAAGAAAGAAAACACTTCTAAGTTCAAGGGCCCA 1545  
 Qy 1378 GGAAGAACTCTTTTCCATTTTGAGAGACCCGACACTCAGCTGAACCCGTACAAAGTGG 1437  
 Db 1546 GCAAGCAATCTGCCCAAGATTAATGAC-----AGCTGAACCTGTTCAGCACTGG 1599  
 Qy 1438 TTGGGAATCTGCAATTAATGCTCATTTTAAGAAAACTAATGAAATGACAGCATCAG- 1496  
 Db 1600 TTGGAGCATGCAATGATGATCAATTTAAATAAGCCGGATACAGCAAGTGTAAAT 1659  
 Qy 1497 ---CCCAACCGGAGCTTCCAGGGCCATCCAGATTTGCAAGAGACATCAAAAATATGCC 1554  
 Db 1660 AACATGGGTGAGATTTCTACAGCCAAAGGCCACTCAAAATGCGCGCCACAAATGTGCT 1719  
 Qy 1555 TGGAGTATACAAAAGTCAAAAAGAACTGATGCTTCTGATTAATGACATCTGTAAAA 1614  
 Db 1720 TGGAGAGATTAAGAGAAAAAATACAGACTCTTAATTA- GAGGTCTTGTCTAAG 1778  
 Qy 1615 CAGCAAAATACATGAATATATGACTGCACTTCAAGTAACCTGAGATTAATCAACAA 1674  
 Db 1779 AATTAATAAATAATCTGTACAGGCGCAAGGCAAGCAAAA----- 1820  
 Qy 1675 GAATGTGTTTGGCTCAGATCTCTTTCTGAACAGAGCAAGACTGAGGTATGAGCCA 1734  
 Db 1821 -----AAGTCAATTTGGGGAGCAAGCAAAATCCAGGGTACCAAGCAG 1863  
 Qy 1735 CCATGGGATTATCAGAAATGATTAAGAGCATTAACATCTCCGTTGGTGTCTCAGG 1794  
 Db 1864 GCATTTGATTC-----AGTACCTGAGATCATCTTTCTCAGCTGAATGAGAAAAG 1917  
 Qy 1795 TTAAAAACAATCAGACAGAAAAACMAAAAGCTGTGTGTGAGCACTTGAATTCAGAGAG 1854  
 Db 1918 TTGAAAACAATCAGACAGAAAAACMAAAGATGAGTGTATCTTGTAGTGTGGAGAA 1977  
 Qy 1855 GTGTGTGTGAGCTTGTAAAGAGATGATGATCTCAAGATATGTGAAGAAGTCTTCAG 1914  
 Db 1978 GTCTGTATGAGATTTCTTAAAGAGCAAAATCTCAAGAAACGTGTAAAGAGTGTAAAG 2037  
 Qy 1915 AATATCAGATGAGGAATAGATCAGATTAATTAATCAAAATGATGTAGAGTTTCTCT 1974  
 Db 2038 AATCTGTGATGAGAAATTTGATTTATGTGTATCAACCAATTAAGAGGAAGTTTCTCT 2097  
 Qy 1975 CTGTGATGAGCAACCTCACTACTGACAAATCAGATGAGTACAGCTTTGAACAATTA 2034  
 Db 2098 CTAGTGAACCACTCATCTCTCTCAGAAAAACAGGCTCAGTTACACTTTGACAGTTTA 2157

Qy 2035 CCAGAAAAATATCGCGAAATATCAATATGCTTCCAGGTTTGTACAGCTTGAAGTCT 2094  
 Db 2158 CAGAAAAATATCTGAAAAAATATCAATATGCTTCCAGAAATTTATAAACCTTGACATCT 2217  
 Qy 2095 AATCTCCCAAAATCACTTATTTTACAGATATGCTAAATGCAATTTGATGAGAAATCT 2154  
 Db 2218 AAAAGCCAAAGGTTATCTACTATACAGATATGCAAGATGATGATTAATGAGAAATAGT 2277  
 Qy 2155 CCGTGTGTGATTTTGAAGTTTGTGTTTATGATGGGTAAATAATACAAACAGAAAGT 2214  
 Db 2278 CCACTGCTGATGTTGAATGATCTGTTTCTATGATGAGCAAAAGATCATAGACTTCTGAT 2337  
 Qy 2215 TTCAATCAGATGATGAAGAAAGACAGGAAGCTTACACTTTTAAAGATGAAGTGAAGT 2274  
 Db 2338 GTATATAGGATTAATGAGAAATCTGGAGATGATACACTGGAAGG---GAGTCGGCTG 2394  
 Qy 2275 AATAGCTTGAAGAGAGATTAATAATGATATGACATGCTAATAGAGGTCAATGAT 2334  
 Db 2395 AGCACTTGTCCGATGAAGTTGATCTTATTTGGATCATGCCAATGAGAGCATTCGTT 2454  
 Qy 2335 TGTTTAGCACTGGAATCTCAATTTTCAAGAGAGAAAGAAACTAGAGTCTCCCTT 2394  
 Db 2455 TGCCTGTCTTGTGAATCTGCAATCAACACAGAGAGAAAGAGTGAATAATTTCTTTG 2514  
 Qy 2395 TTCCCAATATCATAGGAAGAAACCTGGTATGACTAGTTCACTTAAGGCTTATGACCT 2454  
 Db 2515 TTTCATTTACGTTTGTGAGAGAGCCACTTTAGCAAAATCAACAAAGACAGCCAAAC 2574  
 Qy 2455 CCTCTTCTGTGGAATTCAAATTAACCAACAGAGATGAGACTTTTCAACAGAAATGCT 2514  
 Db 2575 CTTTGTGTGATATGCTATGAGAAAGAAAGAAAGAGCAGATTAATGTTAAAGAGATTA 2634  
 Qy 2515 ATGCAATGCTGCTTCTTCAACACAGGACCAATCTTAATCCCTATATGTTAAAT 2574  
 Db 2635 CATGCAAGTCAAGCTTCCACCCAGAGCTTAATCTCAATCCCTCTGATTTCTAT 2694  
 Qy 2575 GAAGACTTGTGCTTACACTACAGCTCTGGAACAGACATCTCTTAATAGTCAAAA 2634  
 Db 2695 GATGTTCTGTATTTCTGCTACACCGTCAACCTTCACTTAATCAATTTCCACAA- 2753  
 Qy 2635 GATGTCTTCTTAATCAGCAACAATTTGAATCTGTTTGTGAAAAATGTGTGTTG 2694  
 Db 2754 -----TACTCAGATATGCTCAAGTCTAAATATGTTTTGTGAAAAATGAGCTGG 2808  
 Qy 2695 GCTACACAGTTAATATGAGAGTGTGTGAGTTCACTTAATGATGGTCCAGTTGTT 2754  
 Db 2809 GCATCTCAGTTAAACAGTGAACAGATAGGTGCTCACTTAATGATGATCTCAACTTGT 2868  
 Qy 2755 GTGACGCGAGAGATCTCTAATCAGTTTATACCTCACAAATGAGTCAAACTAGTAT 2814  
 Db 2869 GTTCAACAGAGGCTCTCTCATATTAATACCGCACAAAGGCAAGATACAGAGCAT 2928  
 Qy 2815 GGAGAAATGAAAAATTAACAGATCAATCAACAGAAATTAACAGTGTCTGTCTTCATC 2874  
 Db 2929 GGAGAGATGACAACTTCCAGATGATCAAAAGACACTGAGTGTCTCTCTCAT 2988  
 Qy 2875 CTTTGAATGTTTCTAATCCGAC 2897  
 Db 2989 CTAAATGCTCTTTGCCAGCTGCTC 3011

RESULT 12  
 AC107053  
 LOCUS AC107053 57057 bp DNA linear PRI 23-MAR-2002  
 DEFINITION Homo sapiens BAC clone RP11-398H1 from 4, complete sequence.  
 ACCESSION AC107053  
 VERSION AC107053.5 GI:19526143  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.



REFERENCE 1 (bases 1 to 57057)  
AUTHORS Sulston,J.B. and Waterston,R.  
TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
MEDLINE 99063792  
PUBMED 9847074

REFERENCE 2 (bases 1 to 57057)  
AUTHORS Isak,A., Kozlowski,A. and Creason,K.  
TITLE The sequence of Homo sapiens BAC clone RP11-398H1  
JOURNAL Unpublished (2001)

REFERENCE 3 (bases 1 to 57057)  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (14-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 57057)  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (21-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 57057)  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (16-MAR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 6 (bases 1 to 57057)  
AUTHORS Waterston,R.  
TITLE Direct Submission  
JOURNAL Submitted (23-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Mar 16, 2002 this sequence version replaced gi:1885145.

----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [sapiens@wustl.wustl.edu](mailto:sapiens@wustl.wustl.edu)  
----- Summary Statistics  
Center project name: H\_NH0398H01  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
The RPc1-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Moon,P.Y., Zhao,B., Frenken,B., Tatenno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries: Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>  
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

FEATURES A single plasmid subclone region exists between bases 27391-27445.  
Source  
1..57057  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="4"  
/map="4"  
/clone="RP11-398H1"  
/clone\_11b="RPc1-11"  
1..127  
repeat\_region  
/rpt\_family="ERV1"  
129..246  
repeat\_region  
/rpt\_family="L1"  
247..532  
repeat\_region  
/rpt\_family="Alu"  
533..587  
repeat\_region  
/rpt\_family="AT-rich"  
761..1075  
repeat\_region  
/rpt\_family="L1"  
1106..1410  
repeat\_region  
/rpt\_family="L2"  
1426..1717  
repeat\_region  
/rpt\_family="Alu"  
1817..2029  
repeat\_region  
/rpt\_family="L1"  
2169..2217  
repeat\_region  
/rpt\_family="U5"  
2316..2359  
repeat\_region  
/rpt\_family="CATRA)n"  
2534..2622  
repeat\_region  
/rpt\_family="ERV1"  
2623..2926  
repeat\_region  
/rpt\_family="Alu"  
2927..3615  
repeat\_region  
/rpt\_family="ERV1"  
3723..4337  
repeat\_region  
/rpt\_family="ERV1"  
4510..4672  
repeat\_region  
/rpt\_family="MALR"  
4736..7003  
repeat\_region  
/rpt\_family="L1"  
7004..7098  
repeat\_region  
/rpt\_family="L1"  
12632..12658  
repeat\_region  
/rpt\_family="AT-rich"  
12659..12998  
repeat\_region  
/rpt\_family="L1"  
13006..13923  
repeat\_region  
/rpt\_family="L1"  
13924..14216  
repeat\_region  
/rpt\_family="Alu"  
14217..15190  
repeat\_region  
/rpt\_family="L1"  
15251..15287  
repeat\_region  
/rpt\_family="(TG)n"  
15363..15631  
repeat\_region  
/rpt\_family="L1"  
15644..15938  
repeat\_region  
/rpt\_family="ERV1"  
15939..15966  
repeat\_region  
/rpt\_family="AT-rich"  
16041..16344  
repeat\_region  
/rpt\_family="ERV1"  
16480..16874  
repeat\_region  
/rpt\_family="L1"  
16884..17266  
repeat\_region  
/rpt\_family="MER1\_type"  
17518..18199  
repeat\_region



```

repeat_region      /rpt_family="L2"
                    18294..18407
repeat_region      /rpt_family="MIR"
                    18703..19006
repeat_region      /rpt_family="Alu"
                    19096..19359
repeat_region      /rpt_family="Alu"
                    19376..19395
repeat_region      /rpt_family="(TAA)n"
                    19416..19715
repeat_region      /rpt_family="Alu"
                    20200..20308
repeat_region      /rpt_family="MALR"
                    20309..21114
repeat_region      /rpt_family="L1"
                    21115..21214
repeat_region      /rpt_family="Alu"
                    21351..21658
repeat_region      /rpt_family="Alu"
                    21676..22978
repeat_region      /rpt_family="L1"
                    23115..23137
repeat_region      /rpt_family="AT_rich"
                    23146..23168
repeat_region      /rpt_family="AT_rich"
                    23421..23513
repeat_region      /rpt_family="L1"
                    23533..23827
repeat_region      /rpt_family="Alu"
                    23833..24130
repeat_region      /rpt_family="Alu"
                    24195..24494
repeat_region      /rpt_family="Alu"
                    24517..24684
repeat_region      /rpt_family="L1"
                    24698..24747
repeat_region      /rpt_family="AT_rich"
                    24787..25165
repeat_region      /rpt_family="L1"
                    25229..25315
repeat_region      /rpt_family="MER2_type"
                    25328..25387
repeat_region      /rpt_family="MER2_type"
                    25426..25733
repeat_region      /rpt_family="Alu"

Query Match      35.2%; Score 1025; DB 9; Length 57057;
Best Local Similarity 98.1%; Pred.No. 8.8e-213;
Matches 1097; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```

```

QY      696 TTTTGTGATAGAGAGCCAGAGACTTATTCACAGTACTTCGTGAGAAATCCACAGA 755
DB      33010 TTTTGTGATAGAGAGCCAGAGACTTATTCACAGTACTTCGTGAGAAATCCACAGA 33069
QY      756 TCGTTAAGTCTGTCTTCAGTAATGGACCATCTTTATATGTCGGAAATCTTCAACAA 815
DB      33070 TCGTTAAGTCTGTCTTCAGTAATGGACCATCTTTATATGTCGGAAATCTTCAACAA 33129
QY      816 AAGTAAAGTTTAGAAGCTGTGGAAGACTCAATATGATATGGGCATGCCAATTTCTAC 875
DB      33130 AAGTAAAGTTTAGAAGCTGTGGAAGACTCAATATGATATGGGCATGCCAATTTCTAC 33189
QY      876 TGCATTTACAGCTCTTCCAGTACAGATTAAGTGTAGTTTATTTGACAAAAGAGACT 935
DB      33190 TGCATTTACAGCTCTTCCAGTACAGATTAAGTGTAGTTTATTTGACAAAAGAGACT 33249
QY      936 TTTGATTTGTGACGCCATCCCAATTAATGAATGATTTCCAAAGATTAAGTTCAAC 995
DB      33250 TTTGATTTGTGACGCCATCCCAATTAATGAATGATTTCCAAAGATTAAGTTCAAC 33309
QY      996 TGAATTTTCTTCTTCAGAGATGGAACGTTTATATCTCACTGTGGGAAATCAAGAAC 1055
DB      33310 TGAATTTTCTTCTTCAGAGATGGAACGTTTATATCTCACTGTGGGAAATCAAGAAC 33369
QY      1056 CAGTATAGTGGAGGGGAAAGATTAATTCAGATGCGAAGAAAGGCCCATTTCTGATA 1115
DB      33370 CAGTATAGTGGAGGGGAAAGATTAATTCAGATGCGAAGAAAGGCCCATTTCTGATA 33429
QY      1116 CCTTGTAGAGCTTATCTCTGATAGATCTGCACTTCTTAATAGTCACTTCAGACAA 1175
DB      33430 CCTTGTAGAGCTTATCTCTGATAGATCTGCACTTCTTAATAGTCACTTCAGACAA 33489
QY      1176 AACATTAATGAGAACATGTCATCTCAGACAGAAATGCTTCAAGTCCAAAAGATCAGG 1235
DB      33490 AACATTAATGAGAACATGTCATCTCAGACAGAAATGCTTCAAGTCCAAAAGATCAGG 33549
QY      1236 AGAGGTGAAATGAAGAGAGGTAATCAACCAAGACAAATAGCCAACTTTTAACTT 1295
DB      33550 AGAGGTGAAATGAAGAGAGGTAATCAACCAAGACAAATAGCCAACTTTTAACTT 33609
QY      1296 CTTTAAAGAAAGACATCCAGTAGTCTTGATCTTTTGAAGACCTGATTAACATCAAGC 1355
DB      33610 CTTTAAAGAAAGACATCCAGTAGTCTTGATCTTTTGAAGACCTGATTAACATCAAGC 33669
QY      1356 ACTCTCAATCACTTTGTCCAGAAAACTCCTTTT 1392
DB      33670 ACTCTCAATCACTTTGTCCAGAAAACTCCTTTT 1392

RESULT 13
AL591646      153342 bp      DNA      linear      HTG 15-AUG-2001
LOCUS      Homo sapiens chromosome 1 clone RP11-25411, 3 unordered pieces.
DEFINITION      AL591646
ACCESSION      AL591646
VERSION      AL591646.6 GI:14586307
KEYWORDS      HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1      Melay,K.
   Direct Submision
   Submitted (14-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
   CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
   Request: clonerequest@sanger.ac.uk
   On Jul 2, 2001 this sequence version replaced gi:14330185.
   ----- Genome Center
   Center: Sanger Centre
   Center code: SC
   Web site: http://www.sanger.ac.uk
   Contact: humquery@sanger.ac.uk

```

----- Project Information  
Center project name: ba25411  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Sequencing vector: plasmid; 108752; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Consensus quality: 153071 bases at least Q40  
Consensus quality: 153095 bases at least Q30  
Insert size: 153142; sum-of-contigs  
Insert size: 154993; 6.8% error; agarose-fp  
Quality coverage: 11.28x in Q20 bases; sum-of-contigs Quality  
coverage: 13.04x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 91998: contig of 91998 bp in length  
\* 91999 92098: gap of 100 bp  
\* 92099 140835: contig of 48737 bp in length  
\* 140836 140935: gap of 100 bp  
\* 140936 153342: contig of 12407 bp in length.  
Location/Qualifiers  
1. .153342  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="RP11-25411"  
/clone\_lib="RP11-1.1"  
1. .91998  
/note="assembly fragment:02569  
fragment chain:1  
clone\_end:T7  
vector\_side:left"  
92099..140835  
/note="assembly fragment:00380  
fragment chain:1"  
140936..153342  
/note="assembly fragment:00644  
fragment chain:1  
clone\_end:Sp6  
vector\_side:right"

## FEATURES

source

misc\_feature  
/note="assembly fragment:02569  
fragment chain:1  
clone\_end:T7  
vector\_side:left"  
92099..140835  
/note="assembly fragment:00380  
fragment chain:1"  
140936..153342  
/note="assembly fragment:00644  
fragment chain:1  
clone\_end:Sp6  
vector\_side:right"

## ORIGIN

Query Match 35.2%; Score 1025; DB 2; Length 153342;  
Best Local Similarity 98.1%; Pred. No. 7.6e-213;  
Matches 1037; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 336 AGCTCGACCTTCATGACCAAGATCATACAGGAGTTGTATCTTCATTCATGATAT 395  
DB 112321 AGCTCGACCTTCATGACCAAGATCATACAGGAGTTGTATCTTCATTCATGATAT 112380  
QY 396 ACTACACCGGAGCTTCACCTTTCTAAGCTTCCTACTGACCTGTAATGAAATCAAGAT 455  
DB 112381 ACTACACCGGAGCTTCACCTTTCTAAGCTTCCTACTGACCTGTAATGAAATCAAGAT 112440  
QY 456 TGGCGATTTGGGGCTGGCACTCAATGAAATCCGATGAAAGACATATACATATG 515  
DB 112441 TGGCGATTTGGGGCTGGCACTCAATGAAATCCGATGAAAGACATATACATATG 112500  
QY 516 TGGAACTCTCACTAATCAATTTACCAAGAAATTCGACTGGAAGTGCATGACCTGTAATC 575  
DB 112501 TGGAACTCTCACTAATCAATTTACCAAGAAATTCGACTGGAAGTGCATGACCTGTAATC 112560  
QY 576 TGATGTTTGGCTCTGGGCTGTATGTTTATACATTAATCTTATGGGAGACCAACCTTCGA 635  
DB 112561 TGATGTTTGGCTCTGGGCTGTATGTTTATACATTAATCTTATGGGAGACCAACCTTCGA 112620

QY 636 CACTGACACAGTCAGAAACACATTAATAAGTAGTATGGCAGATTATGAAATGCCATC 695  
DB 112621 CACTGACACAGTCAGAAACACATTAATAAGTAGTATGGCAGATTATGAAATGCCATC 112680  
QY 696 TTTTGTGCAATPAGAGGCCAAGACCTTATTCACCAAGTTACTTGTGTAAATCCAGCAGA 755  
DB 112681 TTTTGTGCAATPAGAGGCCAAGACCTTATTCACCAAGTTACTTGTGTAAATCCAGCAGA 112740  
QY 756 TCGTTAAGTCGTCTTCAGATTAAGGACCAATCCCTTATATGCCGAATCTTCACAAA 815  
DB 112741 TCGTTAAGTCGTCTTCAGATTAAGGACCAATCCCTTATATGCCGAATCTTCACAAA 112800  
QY 816 AAGTAAAGATTAGGAACCTGTGAAAGACTCAATGATAGTGGGACATGCCAATTTCTAC 875  
DB 112801 AAGTAAAGATTAGGAACCTGTGAAAGACTCAATGATAGTGGGACATGCCAATTTCTAC 112860  
QY 876 TGCATTAACAGCTTCTTCAGTACCAAGTAAAGTGTATTTATTTGCAAAAAGAGCT 935  
DB 112861 TGCATTAACAGCTTCTTCAGTACCAAGTAAAGTGTATTTATTTGCAAAAAGAGCT 112920  
QY 936 TTTGATGTGTCAGCCACTCCCAATTAATGACTGTATTTCCAAAGATTAAGTTCAAC 995  
DB 112921 TTTGATGTGTCAGCCACTCCCAATTAATGACTGTATTTCCAAAGATTAAGTTCAAC 112980  
QY 996 TGATTTTCTTCTTCAGAGAGTGAACAGTTTATTAATCAAGTGGGAAATCAAGAAAC 1055  
DB 112981 TGATTTTCTTCTTCAGAGAGTGAACAGTTTATTAATCAAGTGGGAAATCAAGAAAC 113040  
QY 1056 CAGTAAATGTGGAAGGGGAGAGTAAATTCAGATGCAAGAAAGGCCAATTCCTCGATA 1115  
DB 113041 CAGTAAATGTGGAAGGGGAGAGTAAATTCAGATGCAAGAAAGGCCAATTCCTCGATA 113100  
QY 1116 CTTGTGTAGCTTATTCCTCTGATAGATCTGSCACTTCTAATAGTCAGTCAAGCAA 1175  
DB 113101 CTTGTGTAGCTTATTCCTCTGATAGATCTGSCACTTCTAATAGTCAGTCAAGCAA 113160  
QY 1176 AACATATACATGGAACGATGCTACTCAGCAGAAATGCTTTCAGTGTCCAAAGATCAGG 1235  
DB 113161 AACATATACATGGAACGATGCTACTCAGCAGAAATGCTTTCAGTGTCCAAAGATCAGG 113320  
QY 1236 AGGAGGTGAAATGGAAGAGGTACTCAACCCACAGACAAATGSCCAATTTTAACTT 1295  
DB 113321 AGGAGGTGAAATGGAAGAGGTACTCAACCCACAGACAAATGSCCAATTTTAACTT 113380  
QY 1296 CTTTAAAGAAAGACATCAATGATCTGATCTTTTGAAGACCTGATTAACATCAAGC 1355  
DB 113381 CTTTAAAGAAAGACATCAATGATCTGATCTTTTGAAGACCTGATTAACATCAAGC 113340  
QY 1356 ACTCTCAATCATCTTTGTCCAGAAAACCTCTTTT 1392  
DB 113341 ACTCTCAATCATCTTTGTCCAGAAAACCTCTTTT 113377  
RESULT 14  
AR083570 1600 bp DNA linear PAT 01-SEP-2000  
LOCUS AR083570 Sequence 5 from patent US 5976893.  
DEFINITION AR083570  
ACCESSION AR083570  
VERSION AR083570.1 GI:10010343  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1600)  
AUTHORS Dennis J.W., Hefferman M. and Fode C.  
TITLE Methods for identifying binding partners, agonists, and antagonist  
of a serine/threonine tyrosine kinase  
JOURNAL Patent: US 5976893-A 5 02-NOV-1993;  
FEATURES Location/Qualifiers  
1. .1600  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 33.3%; Score 970.4; DB 6; Length 1600;  
 Best Local Similarity 85.0%; Pred. No. 1.1e-200;  
 Matches 1161; Conservative 0; Mismatches 186; Indels 19; Gaps 6;

1 ATGGCGACCTGCATCGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGGTAA 60  
 206 ATGGCGCGCTGCATCGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGGTAA 265  
 61 GGATCATTTGCTGCTGTCTACAGAGCTGATTCACATCTGCTTGAAGTTGCAATC 120  
 266 GGATCATTTGCTGCTGTCTACAGAGCTGATTCACATCTGCTTGAAGTTGCAATC 325  
 121 AAAATGATGATTAAGAAAGCAATGTACAAAGCAGAAAGTGTACAGAGCTCAAAATGAG 180  
 326 AAAATGATGATTAAGAAAGCAATGTACAAAGCAGAAAGTGTACAGAGCTCAAAATGAG 385  
 181 GTGAAAATACATTTGCCAATTTGAAACATCTTCTATCTGGAGCTTTATTAATTTTGA 240  
 386 GTGAAAATACATTTGCCAATTTGAAACATCTTCTATCTGGAGCTTTATTAATTTTGA 445  
 241 GATAGCAATTAATGTATCTGTATGTAAGAAATGTGCAATATGAGAAATGAACAGATAT 300  
 446 GATTAACAATTAATGTATCTGTATGTAAGAAATGTGCAATATGAGAAATGAACAGATAT 505  
 301 CTAAAGAAATGAGTGAACCTTCTCAGAAAATGAAAGCTCGACATTCATGACCGAGATC 360  
 506 CTGAAGAAACAGATGAAAGCCTTCTCAGAAAATGAAAGCTCGAGATTCATGACCGAGAT 565  
 361 ATCACAGGAGATGTGATCTTCTCATCTGATATCTACACCGGAGCTTCACATCTTCT 420  
 566 ATCACAGGAGATGTGATCTTCTCATCTGATATCTACACCGGAGCTTCACATCTTCT 625  
 421 AACCTCTACTGACTCTGTAATATGAACATCAAGATGCTGATTTGGCTGCGCACTCA 480  
 626 AACATCTTCTACTGACTCTGTAATATGAACATCAAGATGCTGATTTGGCTGCGCACTCA 685  
 481 CTGAAAATGCGCATGAAAAGCATTAATATGGAACCTCTCAATCAATTCACCA 540  
 686 TTGAATATGCGCATGAAAAGCATTAATATGGAACCTCTCAATCAATTCACCA 745  
 541 GAAATGCGCATGAAAGTGCACATGCGCTTGAATCTGATGTTGGCTGCGCTGATG 600  
 746 GAAATGCGCATGAAAGTGCACATGCGCTTGAATCTGATGTTGGCTGCGCTGATG 805  
 601 TTTTATACATTAATTAATCGGAGAGACCACTTCGACATGACACAGTCAAGAACATTA 660  
 806 TCTTATACATTAATTAATCGGAGAGACCACTTCGACATGACACAGTCAAGAACATTA 865  
 661 AATAAAGTATGATGCGCAATTAAGAAATGCGCATCTTTTGTTCATAGAGCGCAAGAC 720  
 866 AATAAAGTATGATGCGCAATTAAGAAATGCGCATCTTTTGTTCATAGAGCGCGCAAGAC 925  
 721 CTTATTCACAGTACTTCTGTAGAAATTCAGAGATGTTTAACTGCTCTTCAGATTTG 780  
 926 CTTATTCACAGTACTTCTGTAGAAATTCAGAGATGTTTAACTGCTCTTCAGATTTG 985  
 781 GACCATCTTTTATGTCGGAATCTTCAACAAAAGATTAAGATTTAGAACTGTGAA 840  
 986 GACCATCTTTTATGTCGGAATCTTCAACAAAAGATTAAGATTTAGAACTGTGAA 1045  
 841 GACTCAATTTGATAGTGGGAGATGCGCAATTTCTACAGCAATTAAGCTTTCCAGTACC 900  
 1046 GACTCAATTTGATAGTGGGAGATGCGCAATTTCTACAGCAATTAAGCTTTCCAGTACC 1105  
 901 AGTATTAAGTATGATTTATTTGCAAAAGAGACTTTGATGGTCAAGCACTCCCAAT 960  
 1106 AGTATTAAGTATGATTTATTTGCAAAAGAGACTTTGATGGTCAAGCACTCCCAAT 1162  
 961 AAAATGATGATTTTCAAAAGATTAAGTCAATGATTTTCTTCTCAGAGATGGA 1020  
 1163 AAAATGATGATTTTCAAAAGATTAAGTCAATGATTTTCTTCTCAGAGATGGA 1219

1021 AACAGTTTATTAATCAATGAGGGAAT-----CAAGAAACAGTAAATGAGAGGGA 1074  
 1220 AGTAATTTTATTAATCAATGAGGGAATTCAGAAACAAAGATTAATGAGGGAAGG 1279  
 1075 AGATTAATTAAGATGAG 1134  
 1280 AGATTAATTAAGATGAG 1339  
 1135 TCTGATTAATTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1194  
 1340 TCTGATTAATTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1396  
 1195 TGTCACTGAG 1254  
 1397 TGTCACTGAG 1453  
 1255 AGTACTGAG 1314  
 1454 AGTACTGAG 1512  
 1315 AGTACTGAG 1360  
 1513 AATGATTAATTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1558

RESULT 15  
 156872 156872 1600 bp DNA 1linear PAT 07-OCT-1997  
 LOCUS Sequence 5 from patent US 5650501.  
 DEFINITION 156872  
 ACCESSION 156872  
 VERSION 156872.1 GI:2477285  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1600)  
 AUTHORS Dennis J.W., Hefferman M. and Fode C.  
 TITLE Serine/threonine kinase and nucleic acid encoding same  
 JOURNAL Patent: US 5650501-A 5 22-JUL-1997;  
 FEATURES  
 source  
 1.1600  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 33.3%; Score 970.4; DB 6; Length 1600;  
 Best Local Similarity 85.0%; Pred. No. 1.1e-200;  
 Matches 1161; Conservative 0; Mismatches 186; Indels 19; Gaps 6;

1 ATGGCGACCTGCATCGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGGTAA 60  
 206 ATGGCGCGCTGCATCGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGGTAA 265  
 61 GGATCATTTGCTGCTGTCTACAGAGCTGATTCACATCTGCTTGAAGTTGCAATC 120  
 266 GGATCATTTGCTGCTGTCTACAGAGCTGATTCACATCTGCTTGAAGTTGCAATC 325  
 121 AAAATGATGATTAAGAAAGCAATGTACAAAGCAGAAAGTGTACAGAGCTCAAAATGAG 180  
 326 AAAATGATGATTAAGAAAGCAATGTACAAAGCAGAAAGTGTACAGAGCTCAAAATGAG 385  
 181 GTGAAAATACATTTGCCAATTTGAAACATCTTCTATCTGGAGCTTTATTAATTTTGA 240  
 386 GTGAAAATACATTTGCCAATTTGAAACATCTTCTATCTGGAGCTTTATTAATTTTGA 445  
 241 GATAGCAATTAATGTATCTGTATGTAAGAAATGTGCAATATGAGAAATGAACAGATAT 300  
 446 GATTAACAATTAATGTATCTGTATGTAAGAAATGTGCAATATGAGAAATGAACAGATAT 505  
 301 CTAAAGAAATGAGTGAACCTTCTCAGAAAATGAAAGCTCGACATTCATGACCGAGATC 360  
 506 CTGAAGAAACAGATGAAAGCCTTCTCAGAAAATGAAAGCTCGAGATTCATGACCGAGAT 565

```
QY 361 ATCAGAGGATGTTGATCTTCAATTCAGGATACACCGGAGCTCACACTTTCT 420
    |||||
Db 566 ATCAGAGGATGTTATATCTTCTCATGCGATATGCAACGGAGCTCACACTCTCT 625
QY 421 AACCTCTTACTGATCTGTAATATGAACATCAAGATTGCTGATTTTGGGCTGGCACTCA 480
    |||||
Db 626 AACATCTTACTTACGGGAATATGAACATTAATAATGCTGACTTGTGACTAGCAACGCA 685
QY 481 CTGAATAATGCCACATGAATAAGCACTATATATGTGAACTCCCTAACATTCACCA 540
    |||||
Db 686 TTGAATATGCCACATGAATAAGCACTATATACCTGTGGAGCTCTAATATATTTCCCA 745
QY 541 GAAATTTGCCACTGAAAGTCACATGAGCTTGAATCTGATGTTTGTCTCTGGGCTGTATG 600
    |||||
Db 746 GAAATTTGCCACTGAAAGTCACATGAGCTTGAATCTGATATTTTGTGACTTGGGCTGTATG 805
QY 601 TTTTATACATTACTTATCGGAGACCACTTTCGACATGACACAGTCAAGAACACATTA 660
    |||||
Db 806 TCTTATACGTTACTTATGGAAGACCACTTTTGCACATGACACAGTCAAGAACACATTA 865
QY 661 AATAAAGTATGATGCGCAATATGAATGCCATCTTTTGTGCAATAGAGCGCAAGAC 720
    |||||
Db 866 AACAAAGTATGCTCTGGCAGATTAAGAAAGCCAGCTTTTGTGACAGAGAGCCGAGAC 925
QY 721 CTTATTCACAGTTACTTGTAGAAATCCAGCAGATCGTTTAACTGTCTTCAATATG 780
    |||||
Db 926 CTTATTCACAGTTACTTGTAGAAATCCGCGAGATCGGTTAAGTCTGTCTTGTGTG 985
QY 781 GACCATCTTTTATATGTCGGAAATCTTCAACAAAAGTAAAGTTAAGAACTGTGAA 840
    |||||
Db 986 GACCATCTTTTATATGTCGGAAATCTTCAACAAAAGTAAAGCTAAGGACTGTAGAG 1045
QY 841 GACTCAATGATATGAGGAGATGCCAATTTCTACTGCAATTAAGCTTCTTCCAGTACC 900
    |||||
Db 1046 GACTCAATGATATGAGGAGATGCCAATTTCTACTGCAATTAAGCTTCTTCTGTATCC 1105
QY 901 AGTATAGTATGATGTTATTTTGAACAAAGAGACTTTGATTTGTGAGCACTCCCAAT 960
    |||||
Db 1106 AGTATAGTATGAGCTTACTTGAC---AGAAAGCTTTGTGTGTCAACCACTTCCAAAT 1162
QY 961 AAAATGCTGATTTTCAAAAGAAATAAGTCACTGATTTTCTTCTTCAAGATGGA 1020
    |||||
Db 1021 AAAATGCTGATTTTCAAAAGAAATAAGTCACTGATTTTCTTCTTCAAGATGGA 1074
    |||||
QY 1220 AGTATTTTGTACTCAATGGGAAATCCAGAAACAAGACTAATATGAGGAGCGGGG 1279
    |||||
Db 1075 AGAGTAATCAAGATGCAAGAAAGGCAATTTCTGATACCTTGTAGAGCTTATTC 1134
    |||||
Db 1280 AGAGTATTTGAAGATGCAAGAAAGGCGGCAATTTCTGATACCTTGTAGAGCTATTC 1339
QY 1135 TCTGATATGATCTGGCACTTCTAATAGTCAAGTCAAGCAAAAACATATACATGAAACGA 1194
    |||||
Db 1340 TCTGATATGAGCCGCCCCCTCTAA---TCAAGTCTGAGCAAAAACATATCTCAGTAGAACGT 1396
QY 1195 TGTCACTCGAGCAAAATGCTTTCAGTGTCCAAAGATCAGAGAGGTGAAAATGAAGAG 1254
    |||||
Db 1397 TGTCACTCGAGCAAAATGCTTTCAAAGCTCAGAAAGATC---ACTGATGAAAATCAACAC 1453
QY 1255 AGGTACTCAACCCACAGCAAAACATGCCACATTTTAACTTTTAAAGAAAAGACATCC 1314
    |||||
Db 1454 AGGTATTTCAACCCACAAAGCAATGTCAA-TGTTTAACTTCATTAACCAACCAAGCC 1512
QY 1315 AGTAGTTCGTGATCTTTTGAAGACCTGATACATCAAGCACTCT 1360
    |||||
Db 1513 AATAGTTAAGATCTTTTGAAGACCGTATATGACTGAGCAAGTAT 1558
    |||||
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 12:53:07 ; Search time 1076 Seconds  
(without alignments)  
11500.930 Million cell updates/sec

Title: US-10-026-021-1

Perfect score: 2913  
Sequence: 1 acggcgaccatgcacgcgggga.....cgaccctatttcattga 2913

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1980s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2911.4	99.9	3937	4	AAK51966 Aak51966 Human pol
2	2911.4	99.9	3937	4	AAI58400 Aai58400 Human pol
3	2911.4	99.9	3937	8	ADB48370 Adb48370 Novel hum
4	2908.2	99.8	3721	8	ACH03903 Ach03903 Human CDN
5	2825.8	97.0	3841	4	AAK52950 Aak52950 Human pol
6	2825.8	97.0	3845	4	AAI60186 Aai60186 Human pol
7	1879	64.5	3447	6	ABI99713 Abi99713 Mouse isc
8	1878	64.5	3447	2	AAT08711 Aat08711 Sak-a ser
9	970.4	33.3	1599	2	AAT08712 Aat08712 Sak-b ser
10	924.2	31.7	1453	2	AAT08710 Aat08710 Sak serin
11	426	14.6	484	8	ACH34348 Ach34348 Human end
12	329.2	11.3	331	6	ABL36507 Abl36507 Human mit
13	284	9.7	609	9	ADD33867 Add33867 Mouse mit
14	282.2	9.7	2585	4	ABL12501 Abl12501 Drosophi1
15	250	8.6	491	8	ACH37505 Ach37505 Human end
16	239	8.2	4961	4	ABL12500 Abl12500 Drosophi1
17	216.8	7.4	302	7	ABT40590 Abt40590 Toxicity
18	154.2	5.3	381	7	ACA56051 Aca56051 Human sig
19	135	4.6	2772	6	ABI99219 Abi99219 Mouse isc
20	134.4	4.6	2770	6	ABQ06831 Abq06831 Disease a
21	134.4	4.6	2783	6	ABQ91962 Abq91962 Human NF-
22	134.4	4.6	2789	2	AAK27228 Aak27228 Human ser
23	134.4	4.6	2976	6	ABQ55069 Abq55069 Human ova

24	133.4	4.6	2781	7	ABT42251 Abt42251 Toxicity
25	133.4	4.6	2781	9	ADB53360 Adb53360 Primary r
26	133.2	4.6	2226	9	ADB59070 Adb59070 Toxicity-
27	133.2	4.6	2226	9	ADB53819 Adb53819 Primary r
28	132.4	4.5	204	4	AA506737 Aa506737 Polynucle
29	132.4	4.5	2783	2	AAK27227 Aak27227 Human ser
30	131.4	4.5	2169	6	ABK84377 Abk84377 Human CDN
31	131.4	4.5	2169	7	ACA56549 Aca56549 Human sig
32	131.4	4.5	2169	8	AD52787 Ad52787 Human PRK
33	131.4	4.5	2169	7	ADA50173 Ada50173 Human pol
34	130.2	4.5	2623	4	ABL11407 Abl11407 Drosophi1
35	124	4.3	2778	4	AAH16619 Aah16619 Human CDN
36	124	4.3	2778	5	AAH78069 Aah78069 Nucleotid
37	123.4	4.2	2348	3	AC83112 Aac83112 DNA encod
38	119	4.1	1485	3	AC46066 Aac46066 Arabidops
39	118	4.1	1209	9	ACF58068 Acf58068 Human HbA
40	118	4.1	2033	6	ABK64870 Abk64870 DNA encod
41	118	4.1	2033	9	ADD89968 Add89968 Human can
42	115.2	4.0	1539	6	AB212357 Ab212357 Arabidops
43	115.2	4.0	1792	3	AC44825 Aac44825 Arabidops
44	115.2	4.0	1825	6	AAI46639 Aai46639 A thalian
45	115.2	4.0	1825	6	AAI46642 Aai46642 A thalian

ALIGNMENTS

RESULT 1	AAK51966	AAK51966 standard; cDNA; 3937 BP.
ID	AAK51966;	
AC	AAK51966;	
XX		
DT	06-NOV-2001	(first entry)
XX		
DE	Human polynucleotide SEQ ID NO 511.	
XX		
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;	
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;	
KW	nervous system disorder; arthritis; inflammation; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	MO200157190-A2.	
XX		
PD		
XX		
XX	09-AUG-2001.	
PF	05-FEB-2001; 2001WC-US004098.	
XX		
PR	03-FEB-2000; 2000US-00496914.	
PR	27-APR-2000; 2000US-00560875.	
PR	20-JUN-2000; 2000US-00598075.	
PR	19-JUL-2000; 2000US-00620325.	
PR	01-SEP-2000; 2000US-00654936.	
PR	15-SEP-2000; 2000US-00663561.	
PR	20-OCT-2000; 2000US-00693325.	
PR	30-NOV-2000; 2000US-00728422.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;	
PI	Ma Y, Zhao Qa, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZM;	
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;	
XX		
XX	WPI; 2001-476283/51.	
DR	P-PSDB; AAM78833.	
XX		
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful	
PT	in diagnosis and gene therapy.	
XX		
PS	Claim 1; Page 1846-1850; 6221pp; English.	

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582), and 3666 (AAW80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication  
XX  
SQ Sequence 3937 BP; 1295 A; 732 C; 773 G; 1137 T; 0 U; 0 Other;

Query Match 99.9%; Score 2911.4; DB 4; Length 3937;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACTGACATCGGGAGAGAGATGAGAGATTTTAAAGTTGGAAATCGCTTGTGAAA 60  
DB 334 ATGGCGACTGACATCGGGAGAGAGATGAGAGATTTTAAAGTTGGAAATCGCTTGTGAAA 393  
QY 61 GGAATCATTTGCTGGTGTCTACAGAGCTGAGTCCATTACACTGGTTTGGAAAGTTGCAATC 120  
DB 394 GGAATCATTTGCTGGTGTCTACAGAGCTGAGTCCATTACACTGGTTTGGAAAGTTGCAATC 453  
QY 121 AAAATGATGATAGAGAAAGCATGTACAAAGCAGAAAGTTGACAGAGATCCCAAAATGAG 180  
DB 454 AAAATGATGATAGAGAAAGCATGTACAAAGCAGAAAGTTGACAGAGATCCCAAAATGAG 513  
QY 181 GTGAAAATACATTGCCAATTGAAAACATCTTCTATCTGGAGCTTATATATCTATTTGAA 240  
DB 514 GTGAAAATACATTGCCAATTGAAAACATCTTCTATCTGGAGCTTATATATCTATTTGAA 573  
QY 241 GATAGCAATTAATGTATCTGATATGAAATGTGCAATATGAGAGAAATGAAACAGTAT 300  
DB 574 GATAGCAATTAATGTATCTGATATGAAATGTGCAATATGAGAGAAATGAAACAGTAT 633  
QY 301 CTAAAGATAGAGTGAACCCCTTCTCAGAAATGAAAGCTCGACATTCATGACAGAGATC 360  
DB 634 CTAAAGATAGAGTGAACCCCTTCTCAGAAATGAAAGCTCGACATTCATGACAGAGATC 693  
QY 361 ATCAGAGGATGTGTATCTTCAATTCATGATATGACACGGGACCTCAACTTTCT 420  
DB 694 ATCAGAGGATGTGTATCTTCAATTCATGATATGACACGGGACCTCAACTTTCT 753  
QY 421 AACCTCTACTGACTCGTATATGAAACATCAAGATTGCGATTTGGGCTGGCACTCAA 480  
DB 754 AACCTCTACTGACTCGTATATGAAACATCAAGATTGCGATTTGGGCTGGCACTCAA 813  
QY 481 CTGAAAATGCGCAGATGAAAAGCATATACATTAATGTGAACTCTTAACTAATTTCA 540  
DB 814 CTGAAAATGCGCAGATGAAAAGCATATACATTAATGTGAACTCTTAACTAATTTCA 873  
QY 541 GAAATTTGCGACATCGAAGTGCACATGGCCTTGAATGTGATGTTTGGTCCCTGGGCTGTATG 600  
DB 874 GAAATTTGCGACATCGAAGTGCACATGGCCTTGAATGTGATGTTTGGTCCCTGGGCTGTATG 933  
QY 601 TTTTATACATTACTTATCGGGAGACCACTTTCGACACTGACAGTCAAGAACACATTA 660  
DB 934 TTTTATACATTACTTATCGGGAGACCACTTTCGACACTGACAGTCAAGAACACATTA 993  
QY 661 AATAAAGATGATTTGGGAGATTAAGAAATGCAATCTTTTGTCAATAGAGCCCAAGAC 720  
DB 994 AATAAAGATGATTTGGGAGATTAAGAAATGCAATCTTTTGTCAATAGAGCCCAAGAC 1053  
QY 721 CTATATCCACAGTACTCTGTAGAAATCAGACAGATGTTTAATCTGTCTTGATATG 780  
DB 1054 CTATATCCACAGTACTCTGTAGAAATCAGACAGATGTTTAATCTGTCTTGATATG 1113  
QY 781 GACCATCTTTATGTGCCGAAATCTTCAACAAAAGTAAAGATTTAGAACCTGTGAA 840

DB 1114 GACCATCTTTATGTGCCGAAATCTTCAACAAAAGTAAAGATTTAGAACCTGTGAA 1173  
QY 841 GACTCAATGATAGTGGGAGATGCCCAATTTCTACATTTACAGCTTCTTCCAGTACC 900  
DB 1174 GACTCAATGATAGTGGGAGATGCCCAATTTCTACATTTACAGCTTCTTCCAGTACC 1233  
QY 901 AGTATAGTGTATGATTTATTTAGCAAAAAGAACCTTTGATTGGTCAAGCACTCCCAAT 960  
DB 1234 AGTATAGTGTATGATTTATTTAGCAAAAAGAACCTTTGATTGGTCAAGCACTCCCAAT 1293  
QY 961 AAAATGACGTATTTTCCAAAGATTAAGTTCACTGATTTTCTTTCAGAGATGGA 1020  
DB 1294 AAAATGACGTATTTTCCAAAGATTAAGTTCACTGATTTTCTTTCAGAGATGGA 1353  
QY 1021 AACAGTTTATATCTCAGTGGGGAATTAAGAAACCAATATATGAGAAAGGGAAGATA 1080  
DB 1354 AACAGTTTATATCTCAGTGGGGAATTAAGAAACCAATATATGAGAAAGGGAAGATA 1413  
QY 1081 ATTCAAGATGCAAGAAAGGCCAATTTCTGATACCTTCTGAGCTTATTTCTCTGAT 1140  
DB 1414 ATTCAAGATGCAAGAAAGGCCAATTTCTGATACCTTCTGAGCTTATTTCTCTGAT 1473  
QY 1141 AGATCTGGCACTTCTTAATATGATCAATCAACAAAACATATATCAATGAAACGATGCAC 1200  
DB 1474 AGATCTGGCACTTCTTAATATGATCAATCAACAAAACATATATCAATGAAACGATGCAC 1533  
QY 1201 TCAGCAAAATGCTTTCAGTGTCCAAAGATCAGAGAGAGTGAATAATGAAAGAGGTAC 1260  
DB 1534 TCAGCAAAATGCTTTCAGTGTCCAAAGATCAGAGAGAGTGAATAATGAAAGAGGTAC 1593  
QY 1261 TCACCCACAGCAACCAATGCCAATTTTAACTTTTAAAGAAAAGACATCCAGTAGT 1320  
DB 1594 TCACCCACAGCAACCAATGCCAATTTTAACTTTTAAAGAAAAGACATCCAGTAGT 1653  
QY 1321 TCTGATCTTTTGAAGAACCTGATTAATCAATCAAGCACTTCCCAATCATCTTTTCCAGGA 1380  
DB 1654 TCTGATCTTTTGAAGAACCTGATTAATCAATCAAGCACTTCCCAATCATCTTTTCCAGGA 1713  
QY 1381 AAAAATCTCTTTTCAATTTGCAAGCCGACACTTCAAGCTGAAACCGTAAACAGTGGTTT 1440  
DB 1714 AAAAATCTCTTTTCAATTTGCAAGCCGACACTTCAAGCTGAAACCGTAAACAGTGGTTT 1773  
QY 1441 GGGAAATCGCAAAATTAATGCTCAATTTAAGAAAACATATGATATGACAGATCAAGCCCA 1500  
DB 1774 GGGAAATCGCAAAATTAATGCTCAATTTAAGAAAACATATGATATGACAGATCAAGCCCA 1833  
QY 1501 AACGGGACCTTCCAGGGCCATCCAGATTGTCAGAGAGACATCAAAAATGCTGAGACT 1560  
DB 1834 AACGGGACCTTCCAGGGCCATCCAGATTGTCAGAGAGACATCAAAAATGCTGAGACT 1893  
QY 1561 GATACAAAAGTCAAAAAGAACTGTATGCTTGTGATATGACATTTCTGTAACAGCAA 1620  
DB 1894 GATACAAAAGTCAAAAAGAACTGTATGCTTGTGATATGACATTTCTGTAACAGCAA 1953  
QY 1621 AATACATGAAATATATGACTGCACTTCAAGTAACTGAGTAAATCCAAAGAAATG 1680  
DB 1954 AATACATGAAATATATGACTGCACTTCAAGTAACTGAGTAAATCCAAAGAAATG 2013  
QY 1681 GTTTTGGCTCAGATCTCTTTTGAACAGAGCAAGACTTGAAGGAGTGAAGCCACCATG 1740  
DB 2014 GTTTTGGCTCAGATCTCTTTTGAACAGAGCAAGACTTGAAGGAGTGAAGCCACCATG 2073  
QY 1741 GGTATCAGAAATGTATTAAGAGCAATTAATCTCCGTTGGTGTCTCAAGGTTAAAA 1800  
DB 2074 GGTATCAGAAATGTATTAAGAGCAATTAATCTCCGTTGGTGTCTCAAGGTTAAAA 2133  
QY 1801 CCAATCAGACAGAAAACCAAAAAGCTGTGTAGACATATCTTGAAGAGAGGTGTGT 1860  
DB 2134 CCAATCAGACAGAAAACCAAAAAGCTGTGTAGACATATCTTGAAGAGAGGTGTGT 2193  
QY 1861 GTGAGCTTGTAAAGAGATGATCTCAAGAAATGTGAAGAAAGTTCTTCAATATCT 1920

```

Db 2194 GTGAGCTTGTAAAGAGATGATGATCTCAAGATATATGTGAAGAGTTCTTCAGATATCT 2253
QY 1921 AGTATGAAATATGATGATCTATTTATTTATTCAAATAGTGTAGAGTTTCTCTTGCT 1980
Db 2254 AGTATGAAATATGATGATCTATTTATTTATTCAAATAGTGTAGAGTTTCTCTTGCT 2313
QY 1981 GATAGACCACTCTCACTCTGACCAATCAGTAGGTACAGCTTTGACATTTTACAGAA 2040
Db 2314 GATAGACCACTCTCACTCTGACCAATCAGTAGGTACAGCTTTGACATTTTACAGAA 2373
QY 2041 AATATCTGCGAAATATCATATGCTTCCAGTTTGTATCAGCTTGTAAATCT 2100
Db 2374 AATATCTGCGAAATATCATATGCTTCCAGTTTGTATCAGCTTGTAAATCT 2433
QY 2101 CCCAAATATCATATTTTCAAGATATGCTTAAATGCAATTTGATGAGAAATCTCTGCT 2160
Db 2434 CCCAAATATCATATTTTCAAGATATGCTTAAATGCAATTTGATGAGAAATCTCTGCT 2493
QY 2161 GCTGATTTTGAAGTTTGTATTTATGATGAGGTTAAATATACAAACAGAAATTCATT 2220
Db 2494 GCTGATTTTGAAGTTTGTATTTATGATGAGGTTAAATATACAAACAGAAATTCATT 2553
QY 2221 CAGGTATTTGAAAGACAGAGAAAGTCTTACCTTTAAAGTGAAGTAAATAC 2280
Db 2554 CAGGTATTTGAAAGACAGAGAAAGTCTTACCTTTAAAGTGAAGTAAATAC 2613
QY 2281 TTGAAAGAGAGATTAATATATATATGATGATGATGATGATGATGATGATGATGAT 2340
Db 2614 TTGAAAGAGAGATTAATATATATATGATGATGATGATGATGATGATGATGATGAT 2673
QY 2341 GCACTGGAATCCATATTTTCAAGAGAGAAAGTGTCTCTCTTTTCCCA 2400
Db 2674 GCACTGGAATCCATATTTTCAAGAGAGAAAGTGTCTCTCTTTTCCCA 2733
QY 2401 ATATATATGAAAGAAAGAACTGTGTATGATGATGATGATGATGATGATGATGATGAT 2460
Db 2734 ATATATATGAAAGAAAGAACTGTGTATGATGATGATGATGATGATGATGATGATGAT 2793
QY 2461 TCTGTGATTTCAATTAATTAACCAACGAGATATGATGATGATGATGATGATGATGAT 2520
Db 2794 TCTGTGATTTCAATTAATTAACCAACGAGATATGATGATGATGATGATGATGATGAT 2853
QY 2521 AGTGTGCTTCTTCCACACAGGACCAATCTTAAATCTCTTATGATGATGATGATGATGAT 2580
Db 2854 AGTGTGCTTCTTCCACACAGGACCAATCTTAAATCTCTTATGATGATGATGATGATGAT 2913
QY 2581 CTTGTGCTTCAACATCAAGCTTCTGGAACAGATCTTCTTAAATGCTTAAAGATGCT 2640
Db 2914 CTTGTGCTTCAACATCAAGCTTCTGGAACAGATCTTCTTAAATGCTTAAAGATGCT 2973
QY 2641 CTTTCTTAATCAGACATCTTTTGAATCTGTTTGTGAAAAATGTTGTTGGCTTACA 2700
Db 2974 CTTTCTTAATCAGACATCTTTTGAATCTGTTTGTGAAAAATGTTGTTGGCTTACA 3033
QY 2701 CAGTTAATAGTGTGAGCTGTGTGAGTTTCAAGTTAATGATGAGGTTCCAGTTGTTGCA 2760
Db 3034 CAGTTAATAGTGTGAGCTGTGTGAGTTTCAAGTTAATGATGAGGTTCCAGTTGTTGCA 3093
QY 2761 GGAGAGTGTCTTCTATCATGATTAATACCTCAACAAATGTCTCAACATGAGTA 2820
Db 3094 GGAGAGTGTCTTCTATCATGATTAATACCTCAACAAATGTCTCAACATGAGTA 3153
QY 2821 AATGAAAAATATACAGATCATCAACAAAGAAATTAACAGTGTCTGCTTCCATCTTTTG 2880
Db 3154 AATGAAAAATATACAGATCATCAACAAAGAAATTAACAGTGTCTGCTTCCATCTTTTG 3213
QY 2881 AATGTTTCTAATCGACTCTATTTTCAATGGA 2913
Db 3214 AATGTTTCTAATCGACTCTATTTTCAATGGA 3246

```

RESULT 2  
AA158400

```

ID AA158400 standard; cDNA; 3937 BP.
XX
AC AA158400;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 603.
XX
KW Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Dmanac RT;
XX
DR WPI, 2001-442253/47.
DR P-PSDB; AAM39244.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 603; 10078bp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AAM38642-AA442213) with nocotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC actinin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 3937 BP; 1295 A; 732 C; 773 G; 1137 T; 0 U; 0 Other;

```

Query Match 99.9%; Score 2911.4; DB 4; Length 3937;  
Best local similarity 100.0%; Pred. No. 0;  
Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGGGCACTGCAATGGGGAAGATCGAGATTTTAAAGTTGAAATCTGTTGGTAA 60
Db 334 ATGGGCACTGCAATGGGGAAGATCGAGATTTTAAAGTTGAAATCTGTTGGTAA 393

```



QY 61 GATCATTTGCTGTGTCTACAGAGCTGATGCTCATCACTGTGTTGGAAAGTTGCATC 120  
Db 394 GATCATTTTGTCTGTGTCTACAGAGCTGATGCTCATCACTGTGTTGGAAAGTTGCATC 453  
QY 121 AAAATGATGATGAAAGGCAATGTAACAAAGCGAATGTTACAGAGCTCCAAAATGAG 180  
Db 454 AAAATGATGATGAAAGGCAATGTAACAAAGCGAATGTTACAGAGCTCCAAAATGAG 513  
QY 181 GTGAAAATACATTTGCCAATGAAACATCTTCTATCTTTGGAGCTTTATATCTATTTTGA 240  
Db 514 GTGAAAATACATTTGCCAATGAAACATCTTCTATCTTTGGAGCTTTATATCTATTTTGA 573  
QY 241 GATAGCAATTAATGTATCTGTATTTGAAATGTGCCATATAGAGAAAATGAAACAGTAT 300  
Db 574 GATAGCAATTAATGTATCTGTATTTGAAATGTGCCATATAGAGAAAATGAAACAGTAT 633  
QY 301 CTAAAGAAATGAGTGAACCCCTTCTGAGAAAATGAAAGCTTGACATTTCAATGACCGAGATC 360  
Db 634 CTAAAGAAATGAGTGAACCCCTTCTGAGAAAATGAAAGCTTGACATTTCAATGACCGAGATC 693  
QY 361 ATCACAGGAGTGTGTATCTTCAATCTCAATGTATCTACACCGGAGCTTCACACTTTCT 420  
Db 694 ATCACAGGAGTGTGTATCTTCAATCTTCAATGTATCTACACCGGAGCTTCACACTTTCT 753  
QY 421 AACCTCTTACTGACTGTATATGAAACATCAAGATTTGATTTGGCTGGCACTCAA 480  
Db 754 AACCTCTTACTGACTGTATATGAAACATCAAGATTTGATTTGGCTGGCACTCAA 813  
QY 481 CTGAAAATGCGCATGAAAGGACATATATATGAGAACTCCATACATCAATTTTCA 540  
Db 814 CTGAAAATGCGCATGAAAGGACATATATATGAGAACTCCATACATCAATTTTCA 873  
QY 541 GAAATTCGCACTGGAAGTGCACATGGCTTGAATCTGAATGTGTGGCTCCCTGGGCTGTATG 600  
Db 874 GAAATTCGCACTGGAAGTGCACATGGCTTGAATCTGAATGTGTGGCTCCCTGGGCTGTATG 933  
QY 601 TTTTATACATTAATCTATCGGGAGACACCGCTTGCAGACTGACACAGTCAAGAACATTA 660  
Db 934 TTTTATACATTAATCTATCGGGAGACACCGCTTGCAGACTGACACAGTCAAGAACATTA 993  
QY 661 AATTAAGTATATGCGCATATATGAAATGCCATCTTTTGTCAATAGAGGCAAGAC 720  
Db 994 AATTAAGTATATGCGCATATATGAAATGCCATCTTTTGTCAATAGAGGCAAGAC 1053  
QY 721 CTATATCCACAGTATCTTGTAGAAATCCAGAGATCGTTTAAGTCTGTCTTCAATATG 780  
Db 1054 CTATATCCACAGTATCTTGTAGAAATCCAGAGATCGTTTAAGTCTGTCTTCAATATG 1113  
QY 781 GACCATCTTTATATGCCCCGGAATCTTCAACAAAAGTAAAGATTAGAACTGTGAA 840  
Db 1114 GACCATCTTTATATGCCCCGGAATCTTCAACAAAAGTAAAGATTAGAACTGTGAA 1173  
QY 841 GACTCAATGATAGTGGGATGCCACAATTTCTACTGCAATTAAGCTTCTTCCAGTACC 900  
Db 1174 GACTCAATGATAGTGGGATGCCACAATTTCTACTGCAATTAAGCTTCTTCCAGTACC 1233  
QY 901 AGTATAGTGTGATTTATTTGACAAAAGAGACTTTGATTTGTCAAGCACTCCCAAT 960  
Db 1234 AGTATAGTGTGATTTATTTGACAAAAGAGACTTTGATTTGTCAAGCACTCCCAAT 1293  
QY 961 AAAATGACGTATTTCCAAAGAAATAAGTTCAACTGATTTTCTTCTTCAGAGATGGA 1020  
Db 1294 AAAATGACGTATTTCCAAAGAAATAAGTTCAACTGATTTTCTTCTTCAGAGATGGA 1353  
QY 1021 AACAGTTTTATATCTCAGTGGGGAATCAAGAAACAGTAATATGGAAGGGAGAGTA 1080  
Db 1354 AACAGTTTTATATCTCAGTGGGGAATCAAGAAACAGTAATATGGAAGGGAGAGTA 1413  
QY 1081 ATTCAAGATGCAAGAAAGGCCACATTTCTGATACCTTCTGAGAGCTTATTTCTCTGAT 1140  
Db 1414 ATTCAAGATGCAAGAAAGGCCACATTTCTGATACCTTCTGAGAGCTTATTTCTCTGAT 1473  
QY 1141 AGATGTGCACTTCAATATAGTCACTCAAGCAAAAACATATCAATGGAACGATGTAC 1200

Db 1474 AGATGTGCACTTCAATATAGTCACTCAAGCAAAAACATATCAATGGAACGATGTAC 1533  
QY 1201 TCAGCAAGAAATCTTCAAGTGTCCAAAAGATCAGAGAGGTGAATAATGAAAGAGTAC 1260  
Db 1534 TCAGCAAGAAATCTTCAAGTGTCCAAAAGATCAGAGAGGTGAATAATGAAAGAGTAC 1593  
QY 1261 TCAGCAAGCAAGCAAGTGGCCAACTTTTAATCTTCTTAAAGAAAGACATCCGATGT 1320  
Db 1594 TCAGCAAGCAAGCAAGTGGCCAACTTTTAATCTTCTTAAAGAAAGACATCCGATGT 1653  
QY 1321 TCTGATCTTTTGAAGAAGCTGATTAACATCAAGCACTCTCCATCATCTTTGTTCAGGA 1380  
Db 1654 TCTGATCTTTTGAAGAAGCTGATTAACATCAAGCACTCTCCATCATCTTTGTTCAGGA 1713  
QY 1381 AAAATCTCTTTTCCATTTGCAAGCCGACACTGACTGAACCGTACAAACAGTGT 1440  
Db 1714 AAAATCTCTTTTCCATTTGCAAGCCGACACTGACTGAACCGTACAAACAGTGT 1773  
QY 1441 GGGAAATCTGCAAAATTAATGCTCATTTTAAGAAAACCTATGAATATGACAGATCAGCCCA 1500  
Db 1774 GGGAAATCTGCAAAATTAATGCTCATTTTAAGAAAACCTATGAATATGACAGATCAGCCCA 1833  
QY 1501 AACCGGACCTTCCAGGGCATTCAGATTTGCAAGAGGACATCAATCAAAAATGCTGAGCT 1560  
Db 1834 AACCGGACCTTCCAGGGCATTCAGATTTGCAAGAGGACATCAATCAAAAATGCTGAGCT 1893  
QY 1561 GATACAAAAGTCAAAAAGAACTCTGATGTCTGTATATGCACTTCTGTAAACAGCA 1620  
Db 1894 GATACAAAAGTCAAAAAGAACTCTGATGTCTGTATATGCACTTCTGTAAACAGCA 1953  
QY 1621 AATACATGAAATATATAGCTGCACTTCAACGTAAACCTGAGATATCCAAAGAAATGT 1680  
Db 1954 AATACATGAAATATATAGCTGCACTTCAACGTAAACCTGAGATATCCAAAGAAATGT 2013  
QY 1681 GTTTTGGCTCAGATCTCTTTCTGAAACAGCAAGACTAGGGGTATGAGCCACATGG 1740  
Db 2014 GTTTTGGCTCAGATCTCTTTCTGAAACAGCAAGACTAGGGGTATGAGCCACATGG 2073  
QY 1741 GGTATACGAATCGTACATTTAAGAAAGATTAATCTCCGTTGGTTGCTCCACAGTTTAAA 1800  
Db 2074 GGTATACGAATCGTACATTTAAGAAAGATTAATCTCCGTTGGTTGCTCCACAGTTTAAA 2133  
QY 1801 CCAATCAGACAGAAAACCAAAAAGGCTGTGGTGAACATCTTGAATTCAGAGAGTGTGT 1860  
Db 2134 CCAATCAGACAGAAAACCAAAAAGGCTGTGGTGAACATCTTGAATTCAGAGAGTGTGT 2193  
QY 1861 GTGAGCTTTGTAAGAGATGATCATTCAGAGATATGTGAAGAAAGTTCTTCAATATCT 1920  
Db 2194 GTGAGCTTTGTAAGAGATGATCATTCAGAGATATGTGAAGAAAGTTCTTCAATATCT 2253  
QY 1921 AGTATGGAATTAAGATCACTATTTATTCCAAATGSGTGAAGGTTTCCCTTGGCT 1980  
Db 2254 AGTATGGAATTAAGATCACTATTTATTCCAAATGSGTGAAGGTTTCCCTTGGCT 2313  
QY 1981 GATAGCACACCTCACCCTACATGACACATCATGATGATAGTATGACAAATTTACAGAA 2040  
Db 2314 GATAGCACACCTCACCCTACATGACACATCATGATGATAGTATGACAAATTTACAGAA 2373  
QY 2041 AAAATCTGGGGAATAATCAATATGCTTCCAGGTTTGTACAGCTTGTGAATCTAAATCT 2100  
Db 2374 AAAATCTGGGGAATAATCAATATGCTTCCAGGTTTGTACAGCTTGTGAATCTAAATCT 2433  
QY 2101 CCCAAATCACTTATTTTCAAGATATGCTAAATGCACTTTTGAATGGAATTCCTCGGT 2160  
Db 2434 CCCAAATCACTTATTTTCAAGATATGCTAAATGCACTTTTGAATGGAATTCCTCGGT 2493  
QY 2161 GCTGATTTTGAAGTTTGGTTTATGATGGGTAAATAACAAAACAGAAAGATTCATT 2220  
Db 2494 GCTGATTTTGAAGTTTGGTTTATGATGGGTAAATAACAAAACAGAAAGATTCATT 2553  
QY 2221 CAGGTATTTGAAAAGACAGGAAAGTCTTACATTTTAAAGTGAAGTAAATATAGC 2280

```
DB 2554 CAGGTGATTGAAAAGCAGGAAAGTCTTACACTTTAAAAAGTGAAGTGAATAC 2613
QY 2281 TTGAAAGAGAGATTAATAATGATATGACCAATGCTAATAGAGGTGATGTTT 2340
DB 2614 TTGAAAGAGAGATTAATAATGATATGACCAATGCTAATAGAGGTGATGTTT 2673
QY 2341 GGACTGGAATCCATATTTTCAAGAGAGAAAGAACTAGAGAGTCCCTTTTCCCA 2400
DB 2674 GCACTGGAATCCATATTTTCAAGAGAGAAAGAACTAGAGAGTCCCTTTTCCCA 2733
QY 2401 AATATCATAGAGAAAGAACTGTAAGTACTAGTTCCTAAGGCTTATCACTCTCT 2460
DB 2734 AATATCATAGAGAAAGAACTGTAAGTACTAGTTCCTAAGGCTTATCACTCTCT 2793
QY 2461 TCTGAGATTCATTAATTAATCCCAAGAGATAGAGATCTTTCAAGAGATGCTAT 2520
DB 2794 TCTGAGATTCATTAATTAATCCCAAGAGATAGAGATCTTTCAAGAGATGCTAT 2853
QY 2521 AGTGTGCTTCTTCCCAAGAGAGACCAATCTTAACTCTATGCTTCAATAGAGAG 2580
DB 2854 AGTGTGCTTCTTCCCAAGAGAGACCAATCTTAACTCTATGCTTCAATAGAGAG 2913
QY 2581 CTGTGCTTCAACTACAGCTTCTGGAAGAGACATCTTCAATAGTCTAAAGATG 2640
DB 2914 CTGTGCTTCAACTACAGCTTCTGGAAGAGACATCTTCAATAGTCTAAAGATG 2973
QY 2641 CTGTGCTTCAACTACAGCTTCTGGAAGAGATCTTCAATAGTCTAAAGATG 2700
DB 2974 CTGTGCTTCAACTACAGCTTCTGGAAGAGATCTTCAATAGTCTAAAGATG 3033
QY 2701 CAGTTAAGTGAAGAGAGTGTGAGTTCAGTTAATAGTGGGCCAGTGGTGGCAG 2760
DB 3034 CAGTTAAGTGAAGAGAGTGTGAGTTCAGTTAATAGTGGGCCAGTGGTGGCAG 3093
QY 2761 GCAGAGAGTGTCTTATCATGATTATACCTGCAAGATGTCACCACTAGTGGT 2820
DB 3094 GCAGAGAGTGTCTTATCATGATTATACCTGCAAGATGTCACCACTAGTGGT 3153
QY 2821 AATGAAATTAACGAGCTACATCAAGAGAAATTCAGTGTCTCTTCCATCTTT 2880
DB 3154 AATGAAATTAACGAGCTACATCAAGAGAAATTCAGTGTCTCTTCCATCTTT 3213
QY 2881 AATGTTTCTAATCCGAGCTCTAATTTTCAATG 2913
DB 3214 AATGTTTCTAATCCGAGCTCTAATTTTCAATG 3246

RESULT 3
ADB48370
ID ADB48370 standard; cDNA, 3937 BP.
XX
AC ADB48370;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel human cDNA SEQ ID NO 280.
XX
KW ss; cancer; neurodegenerative disease; human.
XX
OS Homo sapiens.
XX
PN US2003104529-A1.
XX
PD 05-JUN-2003.
XX
PF 04-JAN-2002; 2002US-00037270.
XX
PR 21-JAN-2000; 2000US-00488725.
XX
PR 25-APR-2000; 2000US-00552317.
XX
PR 19-JUL-2000; 2000US-00620312.
XX
PA (ZHOU/) ZHOU P.
XX
PA (TANG/) TANG Y T.
```

```
PA (LIUC/) LIU C.
PA (ASUNDI/) ASUNDI V.
PA (DRMANAC/) DRMANAC R T.
PI Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;
XX WPI; 2003-678194/64.
XX
PT New polynucleotide, useful for treating diseases e.g., cancer or
PT neurodegenerative diseases.
PS Claim 1; SEQ ID NO 280; 99pp; English.
XX
CC The invention relates to a polynucleotide comprising a sequence given in
CC the specification, or its mature protein-coding portion, or its
CC complement. The polynucleotide is useful for treating diseases e.g.,
CC cancer or neurodegenerative diseases and many others listed in the
CC specification. The present sequence represents a novel human cDNA. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?docID=20030104529.
XX
SQ Sequence 3937 BP; 1295 A; 732 C; 773 G; 1137 T; 0 U; 0 Other;
Query Match 99.9%; Score 2911.4; DB 8; Length 3937;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCACTGTCATCGGGAGAGATCGAGATTTTAAAGTTGAAATCTGCTGTTAA 60
DB 334 ATGGGCACTGTCATCGGGAGAGATCGAGATTTTAAAGTTGAAATCTGCTGTTAA 393
QY 61 GGATCAATTTGCTGCTGCTCAGAGAGTGCATTCACATCGTTTGGAGTTGCAATC 120
DB 394 GGATCAATTTGCTGCTGCTCAGAGAGTGCATTCACATCGTTTGGAGTTGCAATC 453
QY 121 AAAATGATGATTAAGAAAGCCATGTACAAAGAGAGATGTACAGAGAGTCCAAATAG 180
DB 454 AAAATGATGATTAAGAAAGCCATGTACAAAGAGAGATGTACAGAGAGTCCAAATAG 513
QY 181 GTGAAATATCATTCGCAATTTGAACATCTTATCTTGGAGCTTTATACATTTTGA 240
DB 514 GTGAAATATCATTCGCAATTTGAACATCTTATCTTGGAGCTTTATACATTTTGA 573
QY 241 GATAGCAATTAATGATGATGATTAAGAAAGTGCATTAAGAGAAATGAACAGGAT 300
DB 574 GATAGCAATTAATGATGATGATTAAGAAAGTGCATTAAGAGAAATGAACAGGAT 633
QY 301 CTAAAGATAGAGTGAAGCCCTTCTCAGAAATGAAGTGCACACTTCATGACAGATC 360
DB 634 CTAAAGATAGAGTGAAGCCCTTCTCAGAAATGAAGTGCACACTTCATGACAGATC 693
QY 361 ATCAGAGGAGTGTGATCTTCAATCTCAGAGTATCTACACCGGAGCTCACAATTCT 420
DB 694 ATCAGAGGAGTGTGATCTTCAATCTCAGAGTATCTACACCGGAGCTCACAATTCT 753
QY 421 AACCTCTCTGATGCTGATTAATGAACATCAAGATTCGATTTGGGCTGGCACTCAA 480
DB 754 AACCTCTCTGATGCTGATTAATGAACATCAAGATTCGATTTGGGCTGGCACTCAA 813
QY 481 CTGAAATATCCATGAAAGACATATACATTAATGGAACCTCTAATCAATTTCAACA 540
DB 814 CTGAAATATCCATGAAAGACATATACATTAATGGAACCTCTAATCAATTTCAACA 873
QY 541 GAAATTCGACTGAAAGTGCACATGAGCTTGAATCTGATGTTGGTCCCTGGGCTGATG 600
DB 874 GAAATTCGACTGAAAGTGCACATGAGCTTGAATCTGATGTTGGTCCCTGGGCTGATG 933
QY 601 TTTTATACATTAATGAGGAGACCACTTGCACATGACAGACAGTCAAGAACATTA 660
DB 934 TTTTATACATTAATGAGGAGACCACTTGCACATGACAGACAGTCAAGAACATTA 993
QY 661 AATAAAGTATGAGCAGATTAATGAATGCATCTTTTGTCAATAGAGCCAAAGAC 720
```

D	b	994	AAATGAATATATATGACATTAATGAAGCCACTTTTGTGCAATGTAGAGCCAAAGAC	1053
Q	y	721	CTTATTCCACGATTAATCTGTGTAGAAAATCCAGCAGATCGTTTAAGCTGTCTTCAGTATG	780
D	b	1054	CTTATTCCACGATTAATCTGTGTAGAAAATCCAGCAGATCGTTTAAGCTGTCTTCAGTATG	1113
Q	y	781	GACCATCTTTTATGTCCCGAAATCTTCAAAAAAGTAAAGATTAGAACTGTGGAA	840
D	b	1114	GACCATCTTTTATGTCCCGAAATCTTCAAAAAAGTAAAGATTAGAACTGTGGAA	1173
Q	y	841	GACTCAATTGATAGTGGGATGCACAAATTTCTACTGCAATTACAGCTTCCAGTACC	900
D	b	1174	GACTCAATTGATAGTGGGATGCACAAATTTCTACTGCAATTACAGCTTCCAGTACC	1233
Q	y	901	AGATTAAGTGTATGTTTATTGACAAAAGAGACTTTGATTTGTGACGCACTCCCAAT	960
D	b	1234	AGATTAAGTGTATGTTTATTGACAAAAGAGACTTTGATTTGTGACGCACTCCCAAT	1293
Q	y	961	AAAATGACTGTATTTTCCAAAGAAATTAAGTTCAACTGATTTTTTCTTCCAGAGATGGA	1020
D	b	1294	AAAATGACTGTATTTTCCAAAGAAATTAAGTTCAACTGATTTTTTCTTCCAGAGATGGA	1353
Q	y	1021	AACAGTTTTTATATCTCAGTGGGGAAATCAGAAAACAGTAATAGTGGAGGGAGAGTA	1080
D	b	1354	AACAGTTTTTATATCTCAGTGGGGAAATCAGAAAACAGTAATAGTGGAGGGAGAGTA	1413
Q	y	1081	ATTCAAGATGCAAGAAAGGCACTTCTGATTAACCTTGTGAGCTTATTTCTCTGAT	1140
D	b	1414	ATTCAAGATGCAAGAAAGGCACTTCTGATTAACCTTGTGAGCTTATTTCTCTGAT	1473
Q	y	1141	AGATCTGCACTTCTAATAGTCACTCTCAGCAAAAACATATCAATGGAACGATGTGAC	1200
D	b	1474	AGATCTGCACTTCTAATAGTCACTCTCAGCAAAAACATATCAATGGAACGATGTGAC	1533
Q	y	1201	TCAGCAGAAATGCTTTCAGTGTCCAAAAGATCAGAGAGAGTGAATGAGAGAGTAC	1260
D	b	1534	TCAGCAGAAATGCTTTCAGTGTCCAAAAGATCAGAGAGAGTGAATGAGAGAGTAC	1593
Q	y	1261	TCACCCACAGACAACAATGCCACATTTTAACTTCTTAAAGAAAAGATCCAGTATG	1320
D	b	1594	TCACCCACAGACAACAATGCCACATTTTAACTTCTTAAAGAAAAGATCCAGTATG	1653
Q	y	1321	TCGTGATCTTTTGAAGAACCTGATTAACAATCAGACCTGCCAATCATCTTTGTCCAGGA	1380
D	b	1654	TCGTGATCTTTTGAAGAACCTGATTAACAATCAGACCTGCCAATCATCTTTGTCCAGGA	1713
Q	y	1381	AAAACTCTTTTCCATTTGAGACCCGACACCTCAGACTGAAAACGTTACAACAGTGTG	1440
D	b	1714	AAAACTCTTTTCCATTTGAGACCCGACACCTCAGACTGAAAACGTTACAACAGTGTG	1773
Q	y	1441	GGGAATCTGCAAAATATGCTCATTTTAAGAAAACTACTGAATATGACAGCATGCCCA	1500
D	b	1774	GGGAATCTGCAAAATATGCTCATTTTAAGAAAACTACTGAATATGACAGCATGCCCA	1833
Q	y	1501	AACCGGACCTTCCAGGGCCATCCAGATTTGACAGAGAGACACATCCAAAAAATGCCCTGACT	1560
D	b	1834	AACCGGACCTTCCAGGGCCATCCAGATTTGACAGAGAGACACATCCAAAAAATGCCCTGACT	1893
Q	y	1561	GATACAAAAGTCMAAAAAGAACTGTGATGCTTCTGATTAATGACATCTGTGTAACAGCAA	1620
D	b	1894	GATACAAAAGTCMAAAAAGAACTGTGATGCTTCTGATTAATGACATCTGTGTAACAGCAA	1953
Q	y	1621	AATACATGAATATATATGATGCACTTCAAGTAAACCTGAGATTAATCCACAGAGATGT	1680
D	b	1954	AATACATGAATATATATGATGCACTTCAAGTAAACCTGAGATTAATCCACAGAGATGT	2013
Q	y	1681	GTTTTTGGCTCAAGTCTCTTTCTGAAACAGACAGACATGAGGGATATGAGCCCAATGG	1740
D	b	2014	GTTTTTGGCTCAAGTCTCTTTCTGAAACAGACAGACATGAGGGATATGAGCCCAATGG	2073
Q	y	1741	GGTATACGAATGGTCAATTAAGACATTAATCTCCGTTGGTTGCTCACAGGTTAAAA	1800

D	b	2074	GGTATACGAATGGTCAATTAAGACATTAATCAATCTCCGTTGGTTGCTCACAGGTTAAAA	2133
Q	y	1801	CCAAATCAGACAGAAAACCAAAAAGGCTGTGTGAGACATCTTGATTCAGAGAGGTGTGT	1860
D	b	2134	CCAAATCAGACAGAAAACCAAAAAGGCTGTGTGAGACATCTTGATTCAGAGAGGTGTGT	2193
Q	y	1861	GTGAGCTGTGAAAAGAGATGATCTCAAGAAATATGGAAGAAAGTCTTCAATATCT	1920
D	b	2194	GTGAGCTGTGAAAAGAGATGATCTCAAGAAATATGGAAGAAAGTCTTCAATATCT	2253
Q	y	1921	AGTATGGAATATCGATCACTATTTATATCCAAATGCTGTGAGAGTTTCTCTTGCT	1980
D	b	2254	AGTATGGAATATCGATCACTATTTATATATTCAAATGCTGTGAGAGTTTCTCTTGCT	2313
Q	y	1981	GATAGACACCCCTCAGCTACTGACAAACATCAGTATAGTACAGTTGACAAATTTCCGAA	2040
D	b	2314	GATAGACACCCCTCAGCTACTGACAAACATCAGTATAGTACAGTTGACAAATTTCCGAA	2373
Q	y	2041	AAATATCGGCGAAAATATCAATATGCTTCCAGTTTGTACAGTTGTGAATCTAAATCT	2100
D	b	2374	AAATATCGGCGAAAATATCAATATGCTTCCAGTTTGTACAGTTGTGAATCTAAATCT	2433
Q	y	2101	CCGAAAATCACTATTTTATCAGATATGCTAAATGCAATTTGATGAGAAATTTCTCTGT	2160
D	b	2434	CCGAAAATCACTATTTTATCAGATATGCTAAATGCAATTTGATGAGAAATTTCTCTGT	2493
Q	y	2161	GCTGATTTTGAAGTTTGGTTTATGATGGGGTAAAAAATACAAACAGAAAGATTCATT	2220
D	b	2494	GCTGATTTTGAAGTTTGGTTTATGATGGGGTAAAAAATACAAACAGAAAGATTCATT	2553
Q	y	2221	CAGTATGTAAGAAAGACAGGGAAGCTTACCTTTAAAAAGTGAAGTGAATTAATAGC	2280
D	b	2554	CAGTATGTAAGAAAGACAGGGAAGCTTACCTTTAAAAAGTGAAGTGAATTAATAGC	2613
Q	y	2281	TTGAAAAGAGATATAAATATGATATGAGACATGCTATGAGGGTCAATGATTTGTGTA	2340
D	b	2614	TTGAAAAGAGATATAAATATGATATGAGACATGCTATGAGGGTCAATGATTTGTGTA	2673
Q	y	2341	GCACTGGAATCCATTAATTTCAAGAGAGAAAGAAAATGAGAGTGTCTCTTTTCCCA	2400
D	b	2674	GCACTGGAATCCATTAATTTCAAGAGAGAAAGAAAATGAGAGTGTCTCTCTTTTCCCA	2733
Q	y	2401	ATATATCATAGAAAGAAAACCTGATAGTACATGATTAAGGCTTAATCAACCTCCCTCT	2460
D	b	2734	ATATATCATAGAAAGAAAACCTGATAGTACATGATTAAGGCTTAATCAACCTCCCTCT	2793
Q	y	2461	TCTGTGATTCAAATTAATCCCAAGAGATATAGCATCTTTCAACAGAAATGCTATGAT	2520
D	b	2794	TCTGTGATTCAAATTAATCCCAAGAGATATAGCATCTTTCAACAGAAATGCTATGAT	2853
Q	y	2521	AGTGTGCTTCTCCAAACAGGAGCAACAATCCTTAATCCCTATATGTTTCAAAATGAAGA	2580
D	b	2854	AGTGTGCTTCTCCAAACAGGAGCAACAATCCTTAATCCCTATATGTTTCAAAATGAAGA	2913
Q	y	2581	CTTGGTCTTTAACACTACGCTTCTGGAACAGACATCTTCTAATATGCTTAAAGATGT	2640
D	b	2914	CTTGGTCTTTAACACTACGCTTCTGGAACAGACATCTTCTAATATGCTTAAAGATGT	2973
Q	y	2641	CTTCTTAATCAGACAACTTTTGAATCTGTTTTTGTGAAAAATGTGTGGGCTACA	2700
D	b	2974	CTTCTTAATCAGACAACTTTTGAATCTGTTTTTGTGAAAAATGTGTGGGCTACA	3033
Q	y	2701	CAGTTAATCAGAGAGCTGTGTGGGTTCAAGTTTAATGATGGGTCACAGTGTGTGAG	2760
D	b	3034	CAGTTAATCAGAGAGCTGTGTGGGTTCAAGTTTAATGATGGGTCACAGTGTGTGAG	3093
Q	y	2761	GCAGAGATGCTTTTATCACTATTAACCTCAACAAATGTGTCAACAACTAGATAGAGAA	2820
D	b	3094	GCAGAGATGCTTTTATCACTATTAACCTCAACAAATGTGTCAACAACTAGATAGAGAA	3153
Q	y	2821	AATGAAAAATTAACAAGTACATCAACAGAAATTAAGATGTCTGTCTTCAATCTTTTG	2880
D	b	3154	AATGAAAAATTAACAAGTACATCAACAGAAATTAAGATGTCTGTCTTCAATCTTTTG	3213

QY 2881 ATGTTTCTATCCGACTCTTAATTTTCATTTGA 2913  
DB 3214 ATGTTTCTATCCGACTCTTAATTTTCATTTGA 3246

RESULT 4  
ACH03903  
ID ACH03903 standard; cDNA; 3721 BP.  
XX ACH03903;  
XX  
XX 26-SEP-2003 (first entry)  
XX  
XX Human cDNA differentially expressed in lung cancer #108.  
XX  
XX Gene therapy; emphysema; ss; gene; chronic obstructive pulmonary disease;  
XX  
XX respiratory disorder; lung cancer; asthma; human.  
XX  
XX Homo sapiens.  
XX  
XX US2003065157-A1.  
XX  
XX 03-APR-2003.  
XX  
XX 04-APR-2002; 2002US-00116802.  
XX  
XX 04-APR-2001; 2001US-0281593P.  
XX  
XX (LASEK) LASEK A W.  
XX  
XX Lasek AW;  
XX  
XX WPI; 2003-540803/51.  
XX  
XX New combination comprising cDNAs that are differentially expressed in  
XX  
XX respiratory disorder; useful for diagnosing or treating respiratory  
XX  
XX disorders e.g., lung cancer, chronic obstructive pulmonary disease,  
XX  
XX emphysema or asthma.  
XX  
XX Claim 1; Page; 39pp; English.  
XX  
XX The invention relates to a combination comprising cDNAs or their  
XX  
XX complements that are differentially expressed in respiratory disorder.  
XX  
XX The combination is useful for preparing a composition for diagnosing or  
XX  
XX treating respiratory disorders e.g., lung cancer, chronic obstructive  
XX  
XX pulmonary disease, emphysema or asthma. The present sequence represents  
XX  
XX human cDNA differentially expressed during lung cancer  
XX  
XX  
XX Sequence 3721 BP; 1207 A; 702 C; 753 G; 1059 T; 0 U; 0 Other;  
XX

Query Match 99.8%; Score 2908.2; DB 8; Length 3721;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2910; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCGACCTGCATCGGGAGAAAGATCGAGATTTTAAAGTTGGAATCTGCTGTGA 60  
DB 288 ATGGCGACCTGCATCGGGAGAAAGATCGAGATTTTAAAGTTGGAATCTGCTGTGA 347  
QY 61 GGATCATTTTGCTGGTGTCTACAGAGCTGATCATTCACACTGGTTTGGAGTTGCAATC 120  
DB 348 GGATCATTTTGCTGGTGTCTACAGAGCTGATCATTCACACTGGTTTGGAGTTGCAATC 407  
QY 121 AAAATGATGATGAAGAACCATGTACAAGCAGAAATGTAACAGAGTCCAAAATGAG 180  
DB 408 AAAATGATGATGAAGAACCATGTACAAGCAGAAATGTAACAGAGTCCAAAATGAG 467  
QY 181 GTGAAGAAATCATTTGCCAATTTGAAACATCTTCTATCTTGGAGTTTAACTATTTTGA 240  
DB 468 GTGAAGAAATCATTTGCCAATTTGAAACATCTTCTATCTTGGAGTTTAACTATTTTGA 527  
QY 241 GATAGCAATTAATGTATCTGTATTAAGAAATGTGCATTAATGAGAAATGAACAGTAT 300  
DB 241 GATAGCAATTAATGTATCTGTATTAAGAAATGTGCATTAATGAGAAATGAACAGTAT 300

DB 528 GATAGCAATTAATGTATCTGTATTAAGAAATGTGCATTAATGAGAAATGAACAGTAT 587  
QY 301 CTAAAGAAATGATGATAAACCCTTCTCAGAAAATGAAGCTCCGACATTCATGACCAAGATC 360  
DB 588 CTAAAGAAATGATGATAAACCCTTCTCAGAAAATGAAGCTCCGACATTCATGACCAAGATC 647  
QY 361 ATCAGAGGATGTTGATCTTCAATCTCAGATGATATACACCGGGAACCTCAGACTTCT 420  
DB 648 ATCAGAGGATGTTGATCTTCAATCTCAGATGATATACACCGGGAACCTCAGACTTCT 707  
QY 421 AACCTCTATCTGATCTGTAATATGAACATCAAGATTCGTGATTTGGCTGGCAACTCA 480  
DB 708 AACCTCTATCTGATCTGTAATATGAACATCAAGATTCGTGATTTGGCTGGCAACTCA 767  
QY 481 CTGAAGAAATGCAATGAAGAAAGCATATATATATGGAACCTCTACATCACTTTCACA 540  
DB 768 CTGAAGAAATGCAATGAAGAAAGCATATATATGGAACCTCTACATCACTTTCACA 827  
QY 541 GAAATTTGCCACTCGAAGTGCACATGGGCTTGAATCTGATGTTGGTCCCTGGGCTGATG 600  
DB 828 GAAATTTGCCACTCGAAGTGCACATGGGCTTGAATCTGATGTTGGTCCCTGGGCTGATG 887  
QY 601 TTTATATCAATTAATTAATGGAAGACACCTTCGACATGACAGATCAAGAACATTA 660  
DB 888 TTTATATCAATTAATTAATGGAAGACACCTTCGACATGACAGATCAAGAACATTA 947  
QY 661 AATTAAGTATATGAGCAATTAATGGAAGACATCTTTTGTCAATATAGGCCCAAGAC 720  
DB 948 AATTAAGTATATGAGCAATTAATGGAAGACATCTTTTGTCAATATAGGCCCAAGAC 1007  
QY 721 CTTATTTACACAGTTACTTGTGAAATCAGAGATCGTTTAACTGCTGCTCAGATATG 780  
DB 1008 CTTATTTACACAGTTACTTGTGAAATCAGAGATCGTTTAACTGCTGCTCAGATATG 1067  
QY 781 GACCATCTTTTATGTCGGAATTTCTTCAACAAAAGATTAAGAACTGTGAA 840  
DB 1068 GACCATCTTTTATGTCGGAATTTCTTCAACAAAAGATTAAGAACTGTGAA 1127  
QY 841 GACTCAATGATATGAGGAGATCCCAATTTTACTGCAATTAAGCTTCTCCAGTAC 900  
DB 1128 GACTCAATGATATGAGGAGATCCCAATTTTACTGCAATTAAGCTTCTCCAGTAC 1187  
QY 901 AGTATAGGATGATGTTTATTTGACAAAAGAAAGATTTGATTTGTCAGGCCATCCCAAT 960  
DB 1188 AGTATAGGATGATGTTTATTTGACAAAAGAAAGATTTGATTTGTCAGGCCATCCCAAT 1247  
QY 961 AAAATGACTGATTTTCAAAAGATTAAGTTCACTGATTTTCTTCTTCAGAGATGGA 1020  
DB 1248 AAAATGACTGATTTTCAAAAGATTAAGTTCACTGATTTTCTTCTTCAGAGATGGA 1307  
QY 1021 AACAGTTTATATCTCAGTGGGAAATCAAGAACCATGATATGAGAAAGGGAAGATTA 1080  
DB 1308 AACAGTTTATATCTCAGTGGGAAATCAAGAACCATGATATGAGAAAGGGAAGATTA 1367  
QY 1081 AATCAAGATGAGAAAGAAAGGCAATTCGTGATCTTCTGAGAGCTTAATCTCTGAT 1140  
DB 1368 AATCAAGATGAGAAAGAAAGGCAATTCGTGATCTTCTGAGAGCTTAATCTCTGAT 1427  
QY 1141 AGATCTGCACTTCTTAATAGTCACTCAAGCAAAACATATATACATGGAAGATGTCA 1200  
DB 1428 AGATCTGCACTTCTTAATAGTCACTCAAGCAAAACATATATACATGGAAGATGTCA 1487  
QY 1201 TCAGCAGAAATGCTTCAAGTGTCCAAAAGATCAAGAGAGGTGAATGAAGAGGTAC 1260  
DB 1488 TCAGCAGAAATGCTTCAAGTGTCCAAAAGATCAAGAGAGGTGAATGAAGAGGTAC 1547  
QY 1261 TCACCAAGACAGACAAATGCAATTTTAACTTTTAAAGAAAGATCACTCAGTAT 1320  
DB 1548 TCACCAAGACAGACAAATGCAATTTTAACTTTTAAAGAAAGATCACTCAGTAT 1607  
QY 1321 TCTGATCTTTTGAAGACCTGATTAACAATCAAGCACTTCCAAATCATTTTGTCCAGGA 1380  
DB 1608 TCTGATCTTTTGAAGACCTGATTAACAATCAAGCACTTCCAAATCATTTTGTCCAGGA 1667

```
QY 1381 AAAATCCTTTTCATTTGACAGACCCGACACCTGAGACTGAAACCGTACAAAGTGGTT 1440
    |||
Db 1668 AAAATCCTTTTCATTTGACAGACCCGACACCTGAGACTGAAACCGTACAAAGTGGTT 1727
QY 1441 GGGAAATCTGCAGAAATTAATGCTCATTTTAAGAAAACTA CTGAAATATGACAGATCAGCCCA 1500
    |||
Db 1728 GGGAAATCTGCAGAAATTAATGCTCATTTTAAGAAAACTA CTGAAATATGACAGATCAGCCCA 1787
QY 1501 AACCGGAGACTTCCAGGGCCATCCAGATTTGGCAGAAAGGACATCAAAAATATGCTGAGACT 1560
    |||
Db 1788 AACCGGAGACTTCCAGGGCCATCCAGATTTGGCAGAAAGGACATCAAAAATATGCTGAGACT 1847
QY 1561 GATTCAAAAGTCAAAAGAACTCTGATGCTTCTGATATGACATTCGTGTAAGAACAGCAA 1620
    |||
Db 1848 GATTCAAAAGTCAAAAGAACTCTGATGCTTCTGATATGACATTCGTGTAAGAACAGCAA 1907
QY 1621 AATACATGAATATATGATGATGCACTTCAAGTAAACCTTGAGATTAATCCAAAGAAATGT 1680
    |||
Db 1908 AATACATGAATATATGATGATGCACTTCAAGTAAACCTTGAGATTAATCCAAAGAAATGT 1967
QY 1681 GTTTTGGCTCAGATCCTCTTCTGAAACAGAGCAAGACTAGGGGTATGAGGCCACATGG 1740
    |||
Db 1968 GTTTTGGCTCAGATCCTCTTCTGAAACAGAGCAAGACTAGGGGTATGAGGCCACATGG 2027
QY 1741 GGTATTCAGATTCGTACATTAAAGACATTACATCTCCGTTGGTTCCTCAGAGTTAAA 1800
    |||
Db 2028 GGTATTCAGATTCGTACATTAAAGACATTACATCTCCGTTGGTTCCTCAGAGTTAAA 2087
QY 1801 CCAATCAGACAGAAAACCAAAAAGGCTGTGTGAGATACCTGATTCAGAGAGGTGTGT 1860
    |||
Db 2088 CCAATCAGACAGAAAACCAAAAAGGCTGTGTGAGATACCTGATTCAGAGAGGTGTGT 2147
QY 1861 GTGAGACTTGTAAAGAGATATGATCTCAAGATATGTAAAGAAAGTTCTTCAGATATCT 1920
    |||
Db 2148 GTGAGACTTGTAAAGAGATATGATCTCAAGATATGTAAAGAAAGTTCTTCAGATATCT 2207
QY 1921 AGTGATGGAATATACGATCACTATTTATCCAAATGCTGTGAGAGTTTCTCTTGCT 1980
    |||
Db 2208 AGTGATGGAATATACGATCACTATTTATCCAAATGCTGTGAGAGTTTCTCTTGCT 2267
QY 1981 GATAGACACCCCTCAGCTTCTGACACATCAGTAGGTACAGCTTTGACAAATTACAGAA 2040
    |||
Db 2268 GATAGACACCCCTCAGCTTCTGACACATCAGTAGGTACAGCTTTGACAAATTACAGAA 2327
QY 2041 AAATCTGGCGAAAATATTAATATGCTTTCAGAGTTGTATCAGCTTGTAAATCT 2100
    |||
Db 2328 AAATCTGGCGAAAATATTAATATGCTTTCAGAGTTGTATCAGCTTGTAAATCT 2387
QY 2101 CCCAAATCACTTATTTTACAGATATGCTAAATGCACTTTGATGAGAAATTCCTGTGT 2160
    |||
Db 2388 CCCAAATCACTTATTTTACAGATATGCTAAATGCACTTTGATGAGAAATTCCTGTGT 2447
QY 2161 GCTGATTTTGGAGTTTGGTTTATGATGGGTAAATAATACAAAACAGAAATTCATT 2220
    |||
Db 2448 GCTGATTTTGGAGTTTGGTTTATGATGGGTAAATAATACAAAACAGAAATTCATT 2507
QY 2221 CAGGTGATTTGAAAAGACAGGGAAGCTTACACTTTAAAAGTGAAGTAAATATGCG 2280
    |||
Db 2508 CAGGTGATTTGAAAAGACAGGGAAGCTTACACTTTAAAAGTGAAGTAAATATGCG 2567
QY 2281 TTGAAAAGAGAGATAAAATATGATATGACATGCTAATGAGGGTCACTGATTTGTTTA 2340
    |||
Db 2568 TTGAAAAGAGAGATAAAATATGATATGACATGCTAATGAGGGTCACTGATTTGTTTA 2627
QY 2341 GCACCTGGAATCCATTAATTTTCAGAAAGAGAAAGAAAACCTAGAGTCTCCCTTTTCCCA 2400
    |||
Db 2628 GCACCTGGAATCCATTAATTTTCAGAAAGAGAAAGAAAACCTAGAGTCTCCCTTTTCCCA 2687
QY 2401 AATAATCATAGGAAGAAAACCTGTGTACTGATGCTCACTAGAGGCTTATCCTCCCTCT 2460
    |||
Db 2688 AATAATCATAGGAAGAAAACCTGTGTACTGATGCTCACTAGAGGCTTATCCTCCCTCT 2747
```

```
QY 2461 TCTGTGATTTCAATTTACCCAAACGAGATGAGCATCTTTTCAACAGAAATGCTCATGAT 2520
    |||
Db 2748 TCTGTGATTTCAATTTACCCAAACGAGATGAGCATCTTTTCAACAGAAATGCTCATGAT 2807
QY 2521 AGNGCTGCTTCTCCAAACAGGCAACCATCCCTTATCCCTCTGATGATTAACAAATGAGA 2580
    |||
Db 2808 AGNGCTGCTTCTCCAAACAGGCAACCATCCCTTATCCCTCTGATGATTAACAAATGAGA 2867
QY 2581 CTGTGCTTCAACATGACAGCTTCTGAAAACAGACATCTCTTAATATGCTAAAAGATGT 2640
    |||
Db 2868 CTGTGCTTCAACATGACAGCTTCTGAAAACAGACATCTCTTAATATGCTAAAAGATGT 2927
QY 2641 CTTCCTTAATCAGCACTTCTTGAATCTGTTTGGAAAAATGTTGGTGGCTTACA 2700
    |||
Db 2928 CTTCCTTAATCAGCACTTCTTGAATCTGTTTGGAAAAATGTTGGTGGCTTACA 2987
QY 2701 CAGTTAATGATGAGAGCTGTGTGAGTTTCAATGATGATGAGTCCAGTTGGTGTGAG 2760
    |||
Db 2988 CAGTTAATGATGAGAGCTGTGTGAGTTTCAATGATGATGAGTCCAGTTGGTGTGAG 3047
QY 2761 GCAGAGAGTCTTCTTATCAGTTATACCTCACCAATGCTCAACACTAGGTATGAGAA 2820
    |||
Db 3048 GCAGAGAGTCTTCTTATCAGTTATACCTCACCAATGCTCAACACTAGGTATGAGAA 3107
QY 2821 AATGAAAATTTACAGCACTACATCAACAGAAATTAAGTGTCTGCTTCATGCTTTTG 2880
    |||
Db 3108 AATGAAAATTTACAGCACTACATCAACAGAAATTAAGTGTCTGCTTCATGCTTTTG 3167
QY 2881 ATGTTTCTAATCCGACTCCTAATTTTCATTGA 2913
    |||
Db 3168 ATGTTTCTAATCCGACTCCTAATTTTCATTGA 3200

RESULT 5
AAKS2950
ID AAKS2950 standard; cDNA; 3841 BP.
XX
AC AAKS2950;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 2479.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
    vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
    tissue growth factor; immunomodulatory; cancer; leukaemia;
    nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN MO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001MO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
XX
PR 27-APR-2000; 2000US-00560875.
XX
PR 20-JUN-2000; 2000US-00598075.
XX
PR 19-JUL-2000; 2000US-00620325.
XX
PR 01-SEP-2000; 2000US-00654936.
XX
PR 15-SEP-2000; 2000US-00663561.
XX
PR 20-OCT-2000; 2000US-00693325.
XX
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Dymanc RT, Asundi V, Zhou P, Xu C, Cao Y;
    Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
    Jue AJ, Yang Y, Wejhrman T, Goodrich R;
    WPI; 2001-476283/51.
XX
DR P-PSDB; AAM79817.
```

XX Nucleic acid encoding polypeptides with cytokine-like activities, useful  
PT in diagnosis and gene therapy.  
XX

PS Claim 1, Page 4757-4758; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAW78823-AAW80302) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAW80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication  
XX

SO Sequence 3641 BP; 1255 A; 704 C; 764 G; 1120 T; 0 U; 0 Other;

Query Match 97.0%; Score 2825.8; DB 4; Length 3841;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 2901; Conservative 0; Mismatches 12; Indels 8; Gaps 6;

QY 1 ATGGGCACTGTCATCGGGGAGAAATCGAGATTTTAAAGTTGAAATCTGCTGTAA 60  
DB ATGGGCACTGTCATCGGGGAGAAATCGAGATTTTAAAGTTGAAATCTGCTGTAA 287  
QY 61 GGATCATTTGCTGCTGCTCTACAGAGCTGAGTCATTCACAGCTGGTTGGAAAGTTGCAATC 120  
DB GGATCATTTGCTGCTGCTCTACAGAGCTGAGTCATTCACAGCTGGTTGGAAAGTTGCAATC 347  
QY 121 AAAATGATGATGAAGAAGCCATGTA CAAGCAGAAAGTGTACAGAGAGTCCAAAATGAG 180  
DB AAAATGATGATGAAGAAGCCATGTA CAAGCAGAAAGTGTACAGAGAGTCCAAAATGAG 407  
QY 181 GTGAAAATACATTTGCCAATTTGAAACATCTTCTATCTTGAGCTTTATATCTATTTTGA 240  
DB GTGAAAATACATTTGCCAATTTGAAACATCTTCTATCTTGAGCTTTATATCTATTTTGA 467  
QY 408 GTGAAAATACATTTGCCAATTTGAAACATCTTCTATCTTGAGCTTTATATCTATTTTGA 467  
DB GTGAAAATACATTTGCCAATTTGAAACATCTTCTATCTTGAGCTTTATATCTATTTTGA 467  
QY 241 GATGACATTTATGTTGATCTGATATGAAGATGTCATTAATGAGAAATGAAACAGATAT 300  
DB GATGACATTTATGTTGATCTGATATGAAGATGTCATTAATGAGAAATGAAACAGATAT 527  
QY 468 GATGACATTTATGTTGATCTGATATGAAGATGTCATTAATGAGAAATGAAACAGATAT 527  
DB GATGACATTTATGTTGATCTGATATGAAGATGTCATTAATGAGAAATGAAACAGATAT 587  
QY 301 CTAAAGATGATGATGAACCCCTTCTCAGAAATGAAAGTGCACATTCATGACACAGATC 360  
DB CTAAAGATGATGATGAACCCCTTCTCAGAAATGAAAGTGCACATTCATGACACAGATC 587  
QY 528 CTAAAGATGATGATGAACCCCTTCTCAGAAATGAAAGTGCACATTCATGACACAGATC 587  
DB CTAAAGATGATGATGAACCCCTTCTCAGAAATGAAAGTGCACATTCATGACACAGATC 587  
QY 361 ATTCACAGGATGTTGATCTTCTATCTGATATCTACACCGGAGCTCAGACTTTCT 420  
DB ATTCACAGGATGTTGATCTTCTATCTGATATCTACACCGGAGCTCAGACTTTCT 647  
QY 588 ATTCACAGGATGTTGATCTTCTATCTGATATCTACACCGGAGCTCAGACTTTCT 647  
DB ATTCACAGGATGTTGATCTTCTATCTGATATCTACACCGGAGCTCAGACTTTCT 647  
QY 421 AACCTCTCTGATCTGATATGAACATCAAGATGTCATTTGGGCTGGCACTCA 480  
DB AACCTCTCTGATCTGATATGAACATCAAGATGTCATTTGGGCTGGCACTCA 707  
QY 648 AACCTCTCTGATCTGATATGAACATCAAGATGTCATTTGGGCTGGCACTCA 707  
DB AACCTCTCTGATCTGATATGAACATCAAGATGTCATTTGGGCTGGCACTCA 767  
QY 481 CTGAAAATGCAATGAAAAGCACTATACATTAATGGAACCTCTTAATCAATTTGCCA 540  
DB CTGAAAATGCAATGAAAAGCACTATACATTAATGGAACCTCTTAATCAATTTGCCA 767  
QY 541 GAAATGGCACTGCAAGTGCACATGGCTTGAATCTGATGTTGGTCCCTGGCTGTATG 600  
DB GAAATGGCACTGCAAGTGCACATGGCTTGAATCTGATGTTGGTCCCTGGCTGTATG 827  
QY 768 GAAATGGCACTGCAAGTGCACATGGCTTGAATCTGATGTTGGTCCCTGGCTGTATG 827  
DB GAAATGGCACTGCAAGTGCACATGGCTTGAATCTGATGTTGGTCCCTGGCTGTATG 827  
QY 601 TTTTATACATTAATGTTGGGAGACACACCTTGACATGACAGATCAAGAAACATTA 660  
DB TTTTATACATTAATGTTGGGAGACACACCTTGACATGACAGATCAAGAAACATTA 887  
QY 828 TTTTATACATTAATGTTGGGAGACACACCTTGACATGACAGATCAAGAAACATTA 887  
DB TTTTATACATTAATGTTGGGAGACACACCTTGACATGACAGATCAAGAAACATTA 887  
QY 661 AATAAAGATGATGTCAGATTAATGAATGCCATCTTTTGTCAAATAGAGGCAAGAC 720  
DB AATAAAGATGATGTCAGATTAATGAATGCCATCTTTTGTCAAATAGAGGCAAGAC 947  
QY 888 AATAAAGATGATGTCAGATTAATGAATGCCATCTTTTGTCAAATAGAGGCAAGAC 947  
DB AATAAAGATGATGTCAGATTAATGAATGCCATCTTTTGTCAAATAGAGGCAAGAC 947

QY 721 CTTATTCACCAAGTACTTGTGTAAGAAATCCAGAGATGTTTAAGTCTGCTCAGATTTG 780  
DB CTTATTCACCAAGTACTTGTGTAAGAAATCCAGAGATGTTTAAGTCTGCTCAGATTTG 1007  
QY 948 CTTATTCACCAAGTACTTGTGTAAGAAATCCAGAGATGTTTAAGTCTGCTCAGATTTG 1007  
DB CTTATTCACCAAGTACTTGTGTAAGAAATCCAGAGATGTTTAAGTCTGCTCAGATTTG 1007  
QY 781 GACCATCTTTTATGTCGGAAATCTTCAACAAAAGTAAAGATTTAGAACTGTGAA 840  
DB GACCATCTTTTATGTCGGAAATCTTCAACAAAAGTAAAGATTTAGAACTGTGAA 1067  
QY 1008 GACCATCTTTTATGTCGGAAATCTTCAACAAAAGTAAAGATTTAGAACTGTGAA 1067  
DB GACCATCTTTTATGTCGGAAATCTTCAACAAAAGTAAAGATTTAGAACTGTGAA 1067  
QY 841 GACTCAATGATAGTGGGAGATGCAATTTCTACAGCAATTCAGCTTCTTCCAGTAC 900  
DB GACTCAATGATAGTGGGAGATGCAATTTCTACAGCAATTCAGCTTCTTCCAGTAC 1127  
QY 1068 GACTCAATGATAGTGGGAGATGCAATTTCTACAGCAATTCAGCTTCTTCCAGTAC 1127  
DB GACTCAATGATAGTGGGAGATGCAATTTCTACAGCAATTCAGCTTCTTCCAGTAC 1127  
QY 901 AGTATAGTGTAGTATTTATTTGCAAAAAGACCTTTGATTTGTCAGGCACTCCCAAT 960  
DB AGTATAGTGTAGTATTTATTTGCAAAAAGACCTTTGATTTGTCAGGCACTCCCAAT 1187  
QY 1128 AGTATAGTGTAGTATTTATTTGCAAAAAGACCTTTGATTTGTCAGGCACTCCCAAT 1187  
DB AGTATAGTGTAGTATTTATTTGCAAAAAGACCTTTGATTTGTCAGGCACTCCCAAT 1247  
QY 961 AAAATGACTGTATTTTCAAAAGATTAAGTTCACTGATTTTCTTCTCAGAGATGGA 1020  
DB AAAATGACTGTATTTTCAAAAGATTAAGTTCACTGATTTTCTTCTCAGAGATGGA 1247  
QY 1188 AAAATGACTGTATTTTCAAAAGATTAAGTTCACTGATTTTCTTCTCAGAGATGGA 1247  
DB AAAATGACTGTATTTTCAAAAGATTAAGTTCACTGATTTTCTTCTCAGAGATGGA 1247  
QY 1021 AACAGTTTATATCTCAGTGGGAAATCAAGAAACAGTAAATGAGGAGAGATGA 1080  
DB AACAGTTTATATCTCAGTGGGAAATCAAGAAACAGTAAATGAGGAGAGATGA 1307  
QY 1248 AACAGTTTATATCTCAGTGGGAAATCAAGAAACAGTAAATGAGGAGAGATGA 1307  
DB AACAGTTTATATCTCAGTGGGAAATCAAGAAACAGTAAATGAGGAGAGATGA 1307  
QY 1081 AATCAAGATGCAAGAAAGGCACTGATCTGATACCTTCTGATGAGCTTAATCTCTGAT 1140  
DB AATCAAGATGCAAGAAAGGCACTGATCTGATACCTTCTGATGAGCTTAATCTCTGAT 1367  
QY 1308 AATCAAGATGCAAGAAAGGCACTGATCTGATACCTTCTGATGAGCTTAATCTCTGAT 1367  
DB AATCAAGATGCAAGAAAGGCACTGATCTGATACCTTCTGATGAGCTTAATCTCTGAT 1367  
QY 1141 AGATCTGGCACTTCTAATAGTCACTCAAGCAAAATATATACATGAGATGTCAC 1200  
DB AGATCTGGCACTTCTAATAGTCACTCAAGCAAAATATATACATGAGATGTCAC 1427  
QY 1368 AGATCTGGCACTTCTAATAGTCACTCAAGCAAAATATATACATGAGATGTCAC 1427  
DB AGATCTGGCACTTCTAATAGTCACTCAAGCAAAATATATACATGAGATGTCAC 1427  
QY 1201 TCAGCAAAATGCTTCACTGATGTCAAAAGATCAAGAGAGTGAATTAAGAGAGTTC 1260  
DB TCAGCAAAATGCTTCACTGATGTCAAAAGATCAAGAGAGTGAATTAAGAGAGTTC 1487  
QY 1428 TCAGCAAAATGCTTCACTGATGTCAAAAGATCAAGAGAGTGAATTAAGAGAGTTC 1487  
DB TCAGCAAAATGCTTCACTGATGTCAAAAGATCAAGAGAGTGAATTAAGAGAGTTC 1487  
QY 1261 TCAGCAAAATGCTTCACTGATGTCAAAAGATCAAGAGAGTGAATTAAGAGAGTTC 1320  
DB TCAGCAAAATGCTTCACTGATGTCAAAAGATCAAGAGAGTGAATTAAGAGAGTTC 1547  
QY 1488 TCAGCAAAATGCTTCACTGATGTCAAAAGATCAAGAGAGTGAATTAAGAGAGTTC 1547  
DB TCAGCAAAATGCTTCACTGATGTCAAAAGATCAAGAGAGTGAATTAAGAGAGTTC 1547  
QY 1321 TCTGATCTTTTGAAGAGCCTGATTAACATCAAGCACTTCCATCATCTTTGTCCAGGA 1380  
DB TCTGATCTTTTGAAGAGCCTGATTAACATCAAGCACTTCCATCATCTTTGTCCAGGA 1607  
QY 1548 TCTGATCTTTTGAAGAGCCTGATTAACATCAAGCACTTCCATCATCTTTGTCCAGGA 1607  
DB TCTGATCTTTTGAAGAGCCTGATTAACATCAAGCACTTCCATCATCTTTGTCCAGGA 1607  
QY 1381 AAACTCTTTTCCATTGTCAGACCCGACACTCAGACTGAAACCGTACACAGTGTCT 1440  
DB AAACTCTTTTCCATTGTCAGACCCGACACTCAGACTGAAACCGTACACAGTGTCT 1667  
QY 1608 AAACTCTTTTCCATTGTCAGACCCGACACTCAGACTGAAACCGTACACAGTGTCT 1667  
DB AAACTCTTTTCCATTGTCAGACCCGACACTCAGACTGAAACCGTACACAGTGTCT 1667  
QY 1441 GGAATCTGCAAAATTAAGTCTCAATTTAAGAAAACCTAAGATTAAGACAGATCCAGCCCA 1500  
DB GGAATCTGCAAAATTAAGTCTCAATTTAAGAAAACCTAAGATTAAGACAGATCCAGCCCA 1727  
QY 1668 GGAATCTGCAAAATTAAGTCTCAATTTAAGAAAACCTAAGATTAAGACAGATCCAGCCCA 1727  
DB GGAATCTGCAAAATTAAGTCTCAATTTAAGAAAACCTAAGATTAAGACAGATCCAGCCCA 1727  
QY 1501 AACCGGACCTTCCAGGGCCATCAGATTTGCAAGAGAGACATCAAAAATGCTGGAAT 1560  
DB AACCGGACCTTCCAGGGCCATCAGATTTGCAAGAGAGACATCAAAAATGCTGGAAT 1787  
QY 1728 AACCGGACCTTCCAGGGCCATCAGATTTGCAAGAGAGACATCAAAAATGCTGGAAT 1787  
DB AACCGGACCTTCCAGGGCCATCAGATTTGCAAGAGAGACATCAAAAATGCTGGAAT 1787  
QY 1561 GATACAAAAGTCAAAAAGAACTGATGCTTCTGATATGCAATTTCTGTAATAACAGCAA 1620  
DB GATACAAAAGTCAAAAAGAACTGATGCTTCTGATATGCAATTTCTGTAATAACAGCAA 1847  
QY 1788 GATACAAAAGTCAAAAAGAACTGATGCTTCTGATATGCAATTTCTGTAATAACAGCAA 1847  
DB GATACAAAAGTCAAAAAGAACTGATGCTTCTGATATGCAATTTCTGTAATAACAGCAA 1847  
QY 1621 AATACCAAGAAATTAATGATGTCAGACTTCAAGTAAACCTGAAATTAATCAAGAAATG 1680  
DB AATACCAAGAAATTAATGATGTCAGACTTCAAGTAAACCTGAAATTAATCAAGAAATG 1907  
QY 1848 AATACCAAGAAATTAATGATGTCAGACTTCAAGTAAACCTGAAATTAATCAAGAAATG 1907  
DB AATACCAAGAAATTAATGATGTCAGACTTCAAGTAAACCTGAAATTAATCAAGAAATG 1907  
QY 1681 GTTTTGGCTCAGATCTCTTCTGTAAGAGAGCAAGCTAGGGGTATGAGGCAACATG 1740  
DB GTTTTGGCTCAGATCTCTTCTGTAAGAGAGCAAGCTAGGGGTATGAGGCAACATG 1967  
QY 1908 GTTTTGGCTCAGATCTCTTCTGTAAGAGAGCAAGCTAGGGGTATGAGGCAACATG 1967  
DB GTTTTGGCTCAGATCTCTTCTGTAAGAGAGCAAGCTAGGGGTATGAGGCAACATG 1967  
QY 1741 GGTATACGAATGATCAATTAAGAGCAATTAATCAATCTCGTGTGCTCAAGGTTAA 1800  
DB GGTATACGAATGATCAATTAAGAGCAATTAATCAATCTCGTGTGCTCAAGGTTAA 2027  
QY 1968 GGTATACGAATGATCAATTAAGAGCAATTAATCAATCTCGTGTGCTCAAGGTTAA 2027  
DB GGTATACGAATGATCAATTAAGAGCAATTAATCAATCTCGTGTGCTCAAGGTTAA 2027



```

QY 1801 CCAATCAGACAGAAAAACAAAAAGCGTGTGACATCTTATTGATTCAGAGAGGTGCT 1860
DB 2028 CCAATCAGACAGAAAAACAAAAAGCGTGTGACATCTTATTGATTCAGAGAGGTGCT 2087
QY 1861 GTGAGCTTGTAAAGATATGATCTCAAGATATGTGAAAAGAGTTCTTCAGATATCT 1920
DB 2088 GTGAGCTTGTAAAGATATGATCTCAAGATATGTGAAAAGAGTTCTTCAGATATCT 2147
QY 1921 AGTATGAAAATACGATCATATTTATATCAAAATGGT-GGTGAGGTTTTCTCTTGC 1979
DB 2148 AGTATGAAAATACGATCATATTTATATCAAAATGGTGTGAGGTTTTCTCTTGC 2207
QY 1980 T-GATAGACCAACCTCAGCTTAC-TGACAAATCAGTAG--TACAGCTTGCACATTTAC 2035
DB 2208 TGGATAGACCAACCTCAGCTTGCACAAATCAGTAGAGGTGACAGCTTTGACAAATTTAC 2267
QY 2036 CAGAAAAATATCTGCGAAAAATATCAATATGCTTCAGGTTGTATCACTTGTAAATCTA 2095
DB 2268 CAGAAAAATATCTGCGAAAAATATCAATATGCTTCAGGTTGTATCACTTGTAAATCTA 2327
QY 2096 AATCTCCCAAAATCACTTAATTTTACAAGATATGCTAAATGCTTTGATGAGAAATCTC 2155
DB 2328 AATCTCCCAAAATCACTTAATTTTACAAGATATGCTAAATGCTTTGATGAGAAATCTC 2387
QY 2156 CTGAGTGTGATTTGAGGTTGTGTTATATGATGGGTTAAAAATACCAAAACAGAAATTT 2215
DB 2388 CTGAGTGTGATTTGAGGTTGTGTTATATGATGGGTTAAAAATACCAAAACAGAAATTT 2447
QY 2216 TCATTCAGGTGATTTGAAAAGACAGGAGTCTTACCTTTAAAAAGTGAAGTGAAGTTA 2275
DB 2448 TCATTCAGGTGATTTGAAAAGACAGGAGTCTTACCTTTAAAAAGTGAAGTGAAGTTA 2507
QY 2276 ATGAGCTTGAAGAGAGATAAAAATGATATATGACATGCTAAATGAGGTCATCTATTT 2335
DB 2508 ATGAGCTTGAAGAGAGATAAAAATGATATATGACATGCTAAATGAGGTCATCTATTT 2567
QY 2336 GTTTAGACATGGAATCATATATTTGAGAAAGAGAAAGAACTAGAGATGCTCCCTTTT 2395
DB 2568 GTTTAGACATGGAATCATATATTTGAGAAAGAGAAAGAACTAGAGATGCTCCCTTTT 2627
QY 2396 TCCCAATATCATAGAGAAAACTGTGTATGATCTATGTTTACCTTAAGGCTTATACCTC 2455
DB 2628 TCCCAATATCATAGAGAAAACTGTGTATGATCTATGTTTACCTTAAGGCTTATACCTC 2687
QY 2456 CTCCTTCTGTGATCAAAATTAACCAAGAGATAGACATCTTTCAACAGAAATGCTCA 2515
DB 2688 CTCCTTCTGTGATCAAAATTAACCAAGAGATAGACATCTTTCAACAGAAATGCTCA 2747
QY 2516 TGCATATGCTGCTTCCACACAGGACCAATCCTTAATCCCTATATGTTCAAAATG 2575
DB 2748 TGCATATGCTGCTTCCACACAGGACCAATCCTTAATCCCTATATGTTCAAAATG 2807
QY 2576 AAGGACTTGGTCTTCAATACAGCTTCTGGAACAGACATCTTCTAATATGCTAAAG 2635
DB 2808 AAGGACTTGGTCTTCAATACAGCTTCTGGAACAGACATCTTCTAATATGCTAAAG 2867
QY 2636 ATGTCTTCTTAATCAGACAACTTTGAAATCTGTTTGTGAAAATGTGGTGGG 2695
DB 2868 ATGTCTTCTTAATCAGACAACTTTGAAATCTGTTTGTGAAAATGTGGTGGG 2927
QY 2696 CTACAC-AGTTAACTATGAGAGCTGTGTGCTTCAATTAATGATGGTCCAGTTGGT 2754
DB 2928 CTACACAGTTAACTATGAGAGCTGTGTGCTTCAATTAATGATGGTCCAGTTGGT 2987
QY 2755 GTGACGAGAGAGTCTTCTATCAGTATATCTCAACCAAAATGGTCAAA--CAACTAGGT 2812
DB 2988 GTGACGAGAGAGTCTTCTATCAGTATATCTCAACCAAAATGGTCAAAACACTAGGGT 3047
QY 2813 ATGAGAAAAATGAAAAATTAACCAATCAATCAAAAGAAATTAACAGTGTCTGTTCCA 2872
DB 3048 ATGAGAAAAATGAAAAATTAACCAATCAATCAAAAGAAATTAACAGTGTCTGTTCCA 3107
QY 2873 TCCTTTGATGTCTTCTAATCCGACTCCATAATTTCAATGA 2913

```

```

DB 3108 TCCTTTGATGTCTTCTAATCCGACTCCATAATTTCAATGA 3148
RESULT 6
AA160186
ID AA160186 standard; cDNA, 3845 BP.
XX
XX AA160186;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 4175.
XX
XX Human; nootropic; immunosuppressant; cytosolic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokine; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia; ss.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX
XX 26-DEC-2000; 2000WO-US034263.
XX
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-00488725.
XX 25-APR-2000; 2000US-00552317.
XX 20-JUN-2000; 2000US-00598042.
XX 19-JUL-2000; 2000US-00620312.
XX 03-AUG-2000; 2000US-00653450.
XX 14-SEP-2000; 2000US-00662191.
XX 19-OCT-2000; 2000US-00693036.
XX 29-NOV-2000; 2000US-00727344.
XX
XX (HYSB-) HYSBQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Pi Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX Pi Zhou F, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
XX
XX P-PSDB; AAM41030.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 4175; 10078bp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
XX encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytosolic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemia and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
XX Sequence 3845 BP; 1253 A; 712 C; 764 G; 1116 T; 0 U; 0 Other;
XX
XX Query Match 97.0%; Score 2825.8; DB 4; Length 3845;
XX Best Local Similarity 99.3%; Pred. No. 0;

```



Matches 2901; Conservative 0; Mismatches 12; Indels 8; Gaps 6;  
QY 1 ATGGCCACCTGCGATCGGGGAGAGATCGAGATTTTAAAGTTGGAAATCTGCTGGTAA 60  
DB 228 ATGGCCACCTGCGATCGGGGAGAGATCGAGATTTTAAAGTTGGAAATCTGCTGGTAA 287  
QY 61 GGATCATTTTGGTGTCTTACAGAGCTGAGTCCATTACACATGCTTGGAGTTGGAAATC 120  
DB 288 GGATCATTTTGGTGTCTTACAGAGCTGAGTCCATTACACATGCTTGGAGTTGGAAATC 347  
QY 121 AAAATGATGATGAGAAAGCCATGTA CAAGCAGAAAGTGTACAGAGTCCAAAATGAG 180  
DB 348 AAAATGATGATGAGAAAGCCATGTA CAAGCAGAAAGTGTACAGAGTCCAAAATGAG 407  
QY 181 GTGAAAATCATTTGGCCAAATTGAAACATCTTCAATCTGGAGCTTAAATCTATTTTGA 240  
DB 408 GTGAAAATCATTTGGCCAAATTGAAACATCTTCAATCTGGAGCTTAAATCTATTTTGA 467  
QY 241 GATAGCAATTAATGTATCTGTATTAAGAAATGTGCATTAATGAGAAATGAACAGTAT 300  
DB 468 GATAGCAATTAATGTATCTGTATTAAGAAATGTGCATTAATGAGAAATGAACAGTAT 527  
QY 301 CTAAAGAAATAGAGTGAACCCCTTCTCAGAAATGAAGCTGACATTCATGACCAAGATC 360  
DB 528 CTAAAGAAATAGAGTGAACCCCTTCTCAGAAATGAAGCTGACATTCATGACCAAGATC 587  
QY 361 ATCACAGGATGTGTATCTTCAATTCATGATTAATCAACGGGACCTCACAATTTCT 420  
DB 588 ATCACAGGATGTGTATCTTCAATTCATGATTAATCAACGGGACCTCACAATTTCT 647  
QY 421 AACCTCTACTGACTGTATATGAACATCAAGATTTGCTGATTTGGGCTGCGAATCAA 480  
DB 648 AACCTCTACTGACTGTATATGAACATCAAGATTTGCTGATTTGGGCTGCGAATCAA 707  
QY 481 CTGAAAATGCCCATGAAAAAGCACTATATATGTGGAACCTCTAATCAATTCACCA 540  
DB 708 CTGAAAATGCCCATGAAAAAGCACTATATATGTGGAACCTCTAATCAATTCACCA 767  
QY 541 GAAATGGCCACTGGAAGTGCACATGGCCTTGAATCTGAATTTGGTCCCTGGGCTGTATG 600  
DB 768 GAAATGGCCACTGGAAGTGCACATGGCCTTGAATCTGAATTTGGTCCCTGGGCTGTATG 827  
QY 601 TTTTATACATTAATCTTATCGGAGAGACACCTTGCACATGACACAGTCAAGAACATTA 660  
DB 828 TTTTATACATTAATCTTATCGGAGAGACACCTTGCACATGACACAGTCAAGAACATTA 887  
QY 661 AATAAAGTATGTGACAGATTAAGAAATGCCATCTTTTGTCAAATAGAGCCAAAGAC 720  
DB 888 AATAAAGTATGTGACAGATTAAGAAATGCCATCTTTTGTCAAATAGAGCCAAAGAC 947  
QY 721 CTTATTCACCAATTAATCTTGTGAGAAATCCAGCATGCTTTAAGTCTGTCTTCAATTTG 780  
DB 948 CTTATTCACCAATTAATCTTGTGAGAAATCCAGCATGCTTTAAGTCTGTCTTCAATTTG 1007  
QY 781 GACCATCTTTTATGTCCCGAAATCTTCAACAAAATGAAGATTAGGAACGTGGA 840  
DB 1008 GACCATCTTTTATGTCCCGAAATCTTCAACAAAATGAAGATTAGGAACGTGGA 1067  
QY 841 GACTCAATTAATGTGAGGACATGCCAATTTCTACTGCAATTAACAGTTCTTCCAGTACC 900  
DB 1068 GACTCAATTAATGTGAGGACATGCCAATTTCTACTGCAATTAACAGTTCTTCCAGTACC 1127  
QY 901 AGTATTAAGTGTATGTTTATTTGCAAAAGAACTTTTGAATGTGACGCCATCCCAAT 960  
DB 1128 AGTATTAAGTGTATGTTTATTTGCAAAAGAACTTTTGAATGTGACGCCATCCCAAT 1187  
QY 961 AAAATGACTGATTTTCCAAAAGATTAAGTTCAACTGATTTTCTTCTTCCAGAGATGA 1020  
DB 1188 AAAATGACTGATTTTCCAAAAGATTAAGTTCAACTGATTTTCTTCTTCCAGAGATGA 1247  
QY 1021 AACAGTTTTTATCTCAGTGGGGAATCAAGAAACAGTAAATGTGGAAGGGGAAGATTA 1080  
DB 1248 AACAGTTTTTATCTCAGTGGGGAATCAAGAAACAGTAAATGTGGAAGGGGAAGATTA 1307

QY 1081 ATTCAAGATGCAAGAAAGGCCACATTTCTGATACCTTGTGAGAGCTTATTTCTGTAT 1140  
DB 1308 ATTCAAGATGCAAGAAAGGCCACATTTCTGATACCTTGTGAGAGCTTATTTCTGTAT 1367  
QY 1141 AGATCTGGCACTTTTAATAGTACAGTCTCAAGCAAAAATATACATTAAGACATCTAC 1200  
DB 1368 AGATCTGGCACTTTTAATAGTACAGTCTCAAGCAAAAATATACATTAAGACATCTAC 1427  
QY 1201 TCAGCAAAATGCTTCAATGTCCTCAAAATATCAGAGAGAGTGTAAATGAAGAGATAC 1260  
DB 1428 TCAGCAAAATGCTTCAATGTCCTCAAAATATCAGAGAGAGTGTAAATGAAGAGATAC 1487  
QY 1261 TCACCCAGACACAAATGTCCTCAATTTTAATCTTTTAAAGAAAGACATCCAGTATG 1320  
DB 1488 TCACCCAGACACAAATGTCCTCAATTTTAATCTTTTAAAGAAAGACATCCAGTATG 1547  
QY 1321 TCTGATCTTTTGAAGAGCTGATTAACAATCAAGACCTCTCAATCATCTTTGTCCAGGA 1380  
DB 1548 TCTGATCTTTTGAAGAGCTGATTAACAATCAAGACCTCTCAATCATCTTTGTCCAGGA 1607  
QY 1381 AAAATCTCTTTTCCATTTTGCAGACCCGACACCTCAGACCTGAAACCGTAAACAGTGT 1440  
DB 1608 AAAATCTCTTTTCCATTTTGCAGACCCGACACCTCAGACCTGAAACCGTAAACAGTGT 1667  
QY 1441 GGGATCTGCAAAATAATGCTCATTTAAGAAAACTACTGAATATGACAGATCAGCCCA 1500  
DB 1668 GGGATCTGCAAAATAATGCTCATTTAAGAAAACTACTGAATATGACAGATCAGCCCA 1727  
QY 1501 AACCGGAACTTCCAGGGCCATCCAGATTTTGCAGAAAGACACATCAAAAAATGCTGACT 1560  
DB 1728 AACCGGAACTTCCAGGGCCATCCAGATTTTGCAGAAAGACACATCAAAAAATGCTGACT 1787  
QY 1561 GATACAAAAGTCAAAAAGAACTCTGATGCTTCTGATTAATGCAATCTGTAAACAGCAA 1620  
DB 1788 GATACAAAAGTCAAAAAGAACTCTGATGCTTCTGATTAATGCAATCTGTAAACAGCAA 1847  
QY 1621 AATACCATGAATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAAAGAAATGT 1680  
DB 1848 AATACCATGAATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAAAGAAATGT 1907  
QY 1681 GTTTTGGCTCAGATCTCTTTCTGAAACAGACAGACATAGGGGTATGAGCCACATGG 1740  
DB 1908 GTTTTGGCTCAGATCTCTTTCTGAAACAGACAGACATAGGGGTATGAGCCACATGG 1967  
QY 1741 GGTATCAGAAATGTCATTAAGAGACATTAATCTCCGTGTGCTCAGAGTTTAA 1800  
DB 1968 GGTATCAGAAATGTCATTAAGAGACATTAATCTCCGTGTGCTCAGAGTTTAA 2027  
QY 1801 CCAATCAGACAGAAACCAAAAAGGCTGTGTGAGCAATCTGATTCAGAGAGAGTGT 1860  
DB 2028 CCAATCAGACAGAAACCAAAAAGGCTGTGTGAGCAATCTGATTCAGAGAGAGTGT 2087  
QY 1861 GTGAGCTTGTAAAGAGATGATCATCTCAAGAAATATGTAAGAAAGTTCTTCAATATCT 1920  
DB 2088 GTGAGCTTGTAAAGAGATGATCATCTCAAGAAATATGTAAGAAAGTTCTTCAATATCT 2147  
QY 1921 AGTATGAAATTCATCTATTTATTTATCCAAATGTG- GGTAGAGTTTCTCTTGC 1979  
DB 2148 AGTATGAAATTCATCTATTTATTTATCCAAATGTGAGAGTGTGAGGTTTCTCTTGC 2207  
QY 1980 T-GATGACACACCTCAGCTAC-TGACAAATCAGTATAGG-TACAGCTTGCATTTTAC 2035  
DB 2208 TGGATGACACACCTCAGCTACCTTACCAATCAGTATAGGAGTATTTGACATTTTAC 2267  
QY 2036 CAGAAAAATATCGGCAAAATATCAATATGCTTCCAGTTGTGACGTTGTAAAGATCA 2095  
DB 2268 CAGAAAAATATCGGCAAAATATCAATATGCTTCCAGTTGTGACGTTGTAAAGATCA 2327  
QY 2096 AATCTCCAAATCATTAATTTTAAAGATATGCTTAAAGCAATTTGATGAGAAATTC 2155  
DB 2328 AATCTCCAAATCATTAATTTTAAAGATATGCTTAAAGCAATTTGATGAGAAATTC 2387

QY 2156 CTGGTCTGATTTGAGGTTGGTTTATGATGAGGGTAAATACAGAAAACAGAGATT 2215  
 Db 2388 CTGGTCTGATTTGAGGTTGGTTTATGATGAGGGTAAATACAGAAAACAGAGATT 2447  
 QY 2216 TCATTGAGTATGAAAAAGCAGGAAAGTTTCACTTTAAAGTAAAGTGAAGTTA 2275  
 Db 2448 TCATTGAGTATGAAAAAGCAGGAAAGTTTCACTTTAAAGTAAAGTGAAGTTA 2507  
 QY 2276 ATAGCTTGAAGAGAGATTAATAATGTATATGACATGCAATGAGAGTCACTGTTT 2335  
 Db 2508 ATAGCTTGAAGAGAGATTAATAATGTATATGACATGCAATGAGAGTCACTGTTT 2567  
 QY 2336 GTTTAGACATGGAATCCATAATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2395  
 Db 2568 GTTTAGACATGGAATCCATAATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2627  
 QY 2396 TCCCAATATATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2455  
 Db 2628 TCCCAATATATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2687  
 QY 2456 CTCTCTCTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2515  
 Db 2688 CTCTCTCTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2747  
 QY 2516 TGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2575  
 Db 2748 TGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2807  
 QY 2576 AAGGACTTGTCTTACAACTACAGCTTCTGGAACAGACATCTCTTAAATGCTAA 2635  
 Db 2808 AAGGACTTGTCTTACAACTACAGCTTCTGGAACAGACATCTCTTAAATGCTAA 2677  
 QY 2636 ATGTCTCTCTTAAATGAGCAAACTTTTGAATCTGTTTGTGAAAATGTTGTTG 2695  
 Db 2868 ATGTCTCTCTTAAATGAGCAAACTTTTGAATCTGTTTGTGAAAATGTTGTTG 2927  
 QY 2696 CTACAC-AGTTAATCTAGGAGCTGTGAGGTTGAGTTAAATGAGGTTCCAGTTG 2754  
 Db 2928 CTACACAGTTAATCTAGGAGCTGTGAGGTTGAGTTAAATGAGGTTCCAGTTG 2987  
 QY 2755 GTGACGAGCAGAGAGTCTTCTATCACTTATACCTCAACAAATGTCATAA-CA 2812  
 Db 2988 GTGACGAGCAGAGAGTCTTCTATCACTTATACCTCAACAAATGTCATAA-CA 3047  
 QY 2813 ATGAGAGAAATGAAATTAATCAAGACTATCAATCAAAAGAAATTAATCAAGTCT 2872  
 Db 3048 ATGAGAGAAATGAAATTAATCAAGACTATCAATCAAAAGAAATTAATCAAGTCT 3107  
 QY 2873 TCCCTTTGATGTTTCTAATCCGACTCCCTAATTTTCAATTTGA 2913  
 Db 3108 TCCCTTTGATGTTTCTAATCCGACTCCCTAATTTTCAATTTGA 3148

RESULT 7  
 AB19713  
 ID AB19713 standard; cDNA; 3447 BP.  
 AC AB19713;  
 XX  
 XX 07-MAR-2002 (first entry)  
 DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:765.  
 XX  
 XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;  
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.  
 XX  
 OS Mus musculus.  
 XX  
 PN W020018188-A2.  
 XX  
 PD 22-NOV-2001.  
 XX  
 PF 18-MAY-2001; 2001WO-JP004192.

XX 18-MAY-2000; 2000JP-00145977.  
 PR (UNYI-) UNIV NIHON SCHOOL JURIDICAL PERSON.  
 XX  
 PA Iehikawa K, Arai S, Takahashi Y, Nagata T, Ishii Y;  
 PI WPI; 2002-03473/04.  
 DR P-PDB; ABB57273.  
 XX  
 XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
 PT expression levels of particular genes defined in the specification or by  
 PT determining the expression profile of a gene group comprising these  
 PT genes.  
 PS Claim 2; Page 1887-1893; 2690pp; English.  
 XX  
 XX The present invention describes a method for examining ischemic  
 CC conditions, comprising measuring the expression levels of particular  
 CC genes (I) in a test sample or determining the expression profile of a  
 CC gene group in the sample comprising genes selected from (I). The method  
 CC is useful for examining the ischemic condition (e.g. compressive  
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
 CC expression levels of particular genes (AB199202 to AB199912, encoding the  
 CC protein sequences in ABB57020 to ABB57374) or by determining the  
 CC expression profile of a gene group comprising these genes. The expression  
 CC levels or expression profiles produced by these genes are used as an  
 CC indicator when screening for ischemic condition-improving drugs or  
 CC therapeutics for ischemic diseases. AB19913 and AB19914 represent PCR  
 CC primers for a mouse ischaemic condition related sequence, which are used  
 CC in the exemplification of the present invention  
 XX  
 SQ Sequence 3447 BP; 1061 A; 690 C; 754 G; 942 T; 0 U; 0 Other;  
 Query Match 64.5%; Score 1879; DB 6; Length 3447;  
 Best Local Similarity 80.9%; Pred. No. 0;  
 Matches 2362; Conservative 0; Mismatches 410; Indels 147; Gaps 9;  
 QY 1 ATGGCAGCCTGCACTGGGAGAGATGAGAGATTTTAAATGGAATCTGCTTGTAA 60  
 Db 206 ATGGCAGCCTGCACTGGGAGAGATGAGAGATTTTAAATGGAATCTGCTTGTAA 265  
 QY 61 GGATCATTTGCTGTGCTCTACAGAGCTGAGTCACTGATTCACAGTGGTGAATC 120  
 Db 266 GGATCATTTGCTGTGCTCTACAGAGCTGAGTCACTGATTCACAGTGGTGAATC 325  
 QY 121 AAAATGATTAAGAAAGCAGTACAAAGCAGAAATGTAAGAGTCCAAATGAG 180  
 Db 326 AAAATGATTAAGAAAGCAGTACAAAGCAGAAATGTAAGAGTCCAAATGAG 385  
 QY 181 GTGAATAATACATTCGCAATGAAACATCTTCTATCTGAGGCTTATATATTTTGA 240  
 Db 386 GTGAATAATACATTCGCAATGAAACATCTTCTATCTGAGGCTTATATATTTTGA 445  
 QY 241 GATAGCAATTAATGATCTGATTTAGAAATGTCATTAATGAGAAATGAAACAGTAT 300  
 Db 446 GATAGCAATTAATGATCTGATTTAGAAATGTCATTAATGAGAAATGAAACAGTAT 505  
 QY 301 CTAAAGATGAGTGAACCTTCTCAGAAATGAAAGTGCACCTTCACTGACCAATC 360  
 Db 506 CTAAAGATGAGTGAACCTTCTCAGAAATGAAAGTGCACCTTCACTGACCAATC 565  
 QY 361 ATCAGAGGATGTTGATCTTCAATTCATGATTAATCAACGAGGACCTCACTTCT 420  
 Db 566 ATCAGAGGATGTTGATCTTCAATTCATGATTAATCAACGAGGACCTCACTTCT 625  
 QY 421 AACCTCTTCTGATCTGATTAATGAAACATCAAGATGCTGATTTTGGCTGCACTCA 480  
 Db 626 AACCTCTTCTGATCTGATTAATGAAACATCAAGATGCTGATTTTGGCTGCACTCA 685  
 QY 481 CTGAATAATGCAATGAAAGCACTATCATTAATGGAATCCCTTAATCACTTCACTCA 540  
 Db 686 TTGAATATGCCAATGAAAGCACTATCATTAATGGAATCCCTTAATCACTTCACTCA 745

QY 541 GAAATTCGCACTGCAAGTGCACATGCGCTTGAATCTGATGTTTGGTCCCTGGGCTGTAG 600  
DB 746 GAAATTCGCACTGCAAGTGCACATGCGCTTGAATCTGATGTTTGGTCCCTGGGCTGTAG 805  
QY 601 TTTTATACATTAATTATCGGAGACCACTTGCACATGCAAGTCAAGACATTA 660  
DB 806 TCTTATACATTAATTATCGGAGACCACTTGCACATGCAAGTCAAGACATTA 865  
QY 661 AATTAAGTATGATGCGAGATTAATGAATGCCATCTTTTGTCAATGAGGCCAAGAC 720  
DB 866 AACAAAGTGTCTCGGCAATTAATGAATGCCATCTTTTGTCAATGAGGCCAAGAC 925  
QY 721 CTTATTCACAGTCTTGTGTAATTCAGCAGATGCTTAACTGTCTTCAATGAT 780  
DB 926 CTTATTCACAGTCTTGTGTAATTCAGCAGATGCTTAACTGTCTTCAATGAT 985  
QY 781 GACCATCTTTTATGTCGGAAATCTTCAACAAAAGTAAAGATTAGGAATGTGAA 840  
DB 986 GACCATCTTTTATGTCGGAAATCTTCAACAAAAGTAAAGATGAGACTGTAGAG 1045  
QY 841 GACTCAATGATAGTGGGCAATGCGCAATTTCTAATGCAATTAAGCTTCTTCAATACC 900  
DB 1046 GACTCAATGATAGTGGGCAATGCGCAATTTCTAATGCAATTAAGCTTCTTCAATACC 1105  
QY 901 AGTATTAAGTATGATTTATTTGAACAAAAGACCTTTGATGTCAGCCTCCCAAT 960  
DB 1106 AGTATTAAGTATGATTTATTTGAACAAAAGACCTTTGATGTCAGCCTCCCAAT 1162  
QY 961 AAAATGATGATATTTTCAAGAATAAAGTCAATGATTTTCTTCAAGAGATGA 1020  
DB 1163 AAAATGATGATATTTTCAAGAATAAAGTCAATGATTTTCTTCAAGAGATGA 1219  
QY 1021 AACAGTTTATTAATCACTGAGGGGAAT-----CAAGAAACAGTAAATAGTGAAGGGA 1074  
DB 1220 AACAGTTTATTAATCACTGAGGGGAAT-----CAAGAAACAGTAAATAGTGAAGGGA 1279  
QY 1075 AAGATTAATCAAGATGCAAGAAAGGCAATTTCCGATACCTTGTAGAGTTATTC 1134  
DB 1280 AAGATTAATCAAGATGCAAGAAAGGCAATTTCCGATACCTTGTAGAGTTATTC 1339  
QY 1135 TCTGATAGATCTGGCACTTCTAATTAATGCTCAAGCAAAAATATATCAATGAGACA 1194  
DB 1340 TCTGATAGATCTGGCACTTCTAATTAATGCTCAAGCAAAAATATATCAATGAGACA 1396  
QY 1195 TGTCACTCAGCAAGAAATGCTTCAAGTGCMAAAGATCAGAGAGAGTGAAGAG 1254  
DB 1397 TGTCACTCAGTAAATGCTTCAAGGCTTGAAGATC----- 1434  
QY 1255 AGTACTCACCACAGACAAAGTCCAACTTTTAACTTTTAAAGAAAAGACATCC 1314  
DB 1435 ----- 1434  
QY 1315 AGTATGTTCTGATCTTTGAAAAGCTGATTAACATCAAGCACTCTCAATCATCTTGT 1374  
DB 1435 -----ACTGATGAAAATCAACACAGTTCATCATCTTGT 1471  
QY 1375 CCAGAAAAATCTCTTTTCATTTGCAAGCCGCACTTCAAGTGAACCGTAAACAG 1434  
DB 1472 CTAGAAAAATCTCTTTTCATTTGCAAGCCGCACTTCAAGTGAACCGTAAACAG 1531  
QY 1435 TGTGTTGGGAATCTGMAATGAAATGCTCAATTTAGGAAAATATATAGCAACCAACGTT 1591  
DB 1532 TGTGTTGGGAATCTGMAATGAAATGCTCAATTTAGGAAAATATATAGCAACCAACGTT 1591  
QY 1495 AGCCCAACCGGGACTTCCAGGGCATCCAGATTTTCAAGAGACATCAAAAATATGCC 1554  
DB 1592 AGCCCAACCGGGAATTTCCAGGGCATCCAGATTTTCCAGATTTTCCAGATTTTCCAG 1648  
QY 1555 TGGACTGATACAAAAGTCAAAAAGAACTGTGATGCTTGTGATATGCAATCTGTAA 1614  
DB 1649 TGGACTGATACAAAAGTCAAAAAGAACTGTGATGCTTGTGATATGCAATCTGTAA 1708

QY 1615 CAGCAAAATACATGAAATATATGACTGCACTTCAAGTAAACCTGATAAATCCACAA 1674  
DB 1709 CAGCTGAGTGCATGAAATATATGAGTGCATCTCAACATTAACCTGAGGCTATGCCAG 1768  
QY 1675 GAATGTTTTTGGCTCAGATCTCTTTCTGAAACAGAGCAAGACTAGGGGATATGAGCA 1734  
DB 1769 GA-----GCCGGGCTCATCTCTCAATCTGAACAAAGCAAAATGAAATGAGTGTG 1822  
QY 1735 CCATGGGTTATCAGAAATGTCATTTAAGAACATTAATCTCCGTTGGTGTCTCAG 1794  
DB 1823 ACATGGGTTATCAGAAATGTCATTTAAGAACATTAATCTCCGTTGGTGTCTCAG 1882  
QY 1795 TTTAAACCAATCAGACAGAAAACAAAAGCTGTGTGTGAGATCTTGAATCAGAGAG 1854  
DB 1883 TTTAAACCAATCAGACAGAAAACAAAAGCTGTGTGTGAGATCTTGAATCAGAGAG 1942  
QY 1855 GTGTGTGTGAGCTTGTAAAGAGATGATCTTCAAGATATGTGAAGAGTTCTTCA 1914  
DB 1943 GTGTGTGTGAGCTTGTGAAGAGATGATCTTCAAGATATGTGAAGAGTTCTTCA 2002  
QY 1915 AATATGATGATGAAATTCATCTAATTTATTTATCCAAATGATGATGATTTTCT 1974  
DB 2003 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2062  
QY 1975 CTTGCTGATGACCAACCTCTCACTGACCAATCAGATGATGATGATGATGATGATGAT 2034  
DB 2063 CTTGCTGATGACCAACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2122  
QY 2035 CCAGAAAAATTAATGAGGAAATATCAATATGCTTCCAGGTTGTGACAGTTGATGAT 2094  
DB 2123 CCAGAAAAATTAATGAGGAAATATCAATATGCTTCCAGGTTGTGATGATGATGAT 2182  
QY 2095 AATATCCCAAAATCACTTATTTTCAAGATATGCTTAAATGCAATTTGATGAGAAAT 2154  
DB 2183 AATATCCCAAAATCACTTATTTTCAAGATATGCTTAAATGCAATTTGATGAGAAAT 2242  
QY 2155 CTTGCTGATGATTTTGAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGAT 2214  
DB 2243 CTTGCTGATGATTTTGAAGGTTTGTGATGATGATGATGATGATGATGATGATGATGAT 2302  
QY 2215 TTCAATCAGTATGATGAAAGACAGGAGTCTTCACTTTTAAAGTGAAGTGAAT 2274  
DB 2303 TTCAATCAGTATGATGAAAGACAGGAGTCTTCACTTTTAAAGTGAAGTGAAT 2362  
QY 2275 AATAGTTAAAGAGAGATTAATAATGATATGAGCAATGCTTAAAGAGGCTATGAT 2334  
DB 2363 AATAGTTAAAGAGAGATTAATAATGATATGAGCAATGCTTAAAGAGGCTATGAT 2422  
QY 2335 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2394  
DB 2423 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2482  
QY 2395 TTCCCAATATCATGAGAAAGAAACCTGTGATGATGATGATGATGATGATGATGAT 2454  
DB 2483 TTCCCAATATCATGAGAAAGAAACCTGTGATGATGATGATGATGATGATGATGAT 2542  
QY 2455 CTTCTCTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2514  
DB 2543 CTTCTCTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2599  
QY 2515 ATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2574  
DB 2600 ATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2659  
QY 2575 GAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2634  
DB 2660 GAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2708  
QY 2635 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2694  
DB 2709 -----GTTCTTCTTAAATCTGCAAGCTTTTGAATCTGTTTGTGAATAATGTTGTG 2764  
QY 2695 GCTACACAGTAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2754

Db 2765 GCTACACAGCTTAAGTACGGAGCTGTGTGGTTCACTTAAATGAGTGGTCACTGGTT 2824  
QY 2755 GTGAGCAGCAGAGCTCTTCTATAGTTAACTCACTCAACAAATGCTCAACACTAGTAT 2814  
Db 2825 GTCCAGCAGAGAGTATCTTCATCAGTTACATCACAGATGCTCAGACACTAGTAT 2884  
QY 2815 GGAAGAAATGAAATATACCACTACATCAACAGAAATTAAGTCTGTCTTCATC 2874  
Db 2885 GGAAGAAATGAAATATACCTGAATCATCAACAGAAATTAAGTCTGTCTTCATC 2944  
QY 2875 CTTTGTATGTTTCTTAATCCGACTCTTAATTTTCATGA 2913  
Db 2945 CTTCTGATGTTTCTTAATCCACTCTTAATTTTCATGA 2983

RESULT 8  
AAT08711  
ID AAT08711 standard; cDNA; 3447 BP.  
AC AAT08711;  
XX  
XX 25-MAY-1996 (first entry)  
XX  
DE Sak-a serine-threonine kinase cDNA.  
XX  
XX Sak-a; serine-threonine kinase; agonist; antagonist;  
XX proliferative disease; cancer; tumour; antisense; transgenic animal;  
XX therapy; ss.  
XX  
XX Mus musculus.  
XX  
XX  
XX Key Location/Qualifiers  
XX CDS 206..2983  
XX /tag= a  
XX /note= "T bases may also be U"  
XX misc\_difference 2767  
XX /tag= b  
XX /note= "base n at position 2767 is unclear in the  
XX specification; encoded amino acid is Ala"  
XX  
XX CA2150789-A.  
XX  
XX 03-DEC-1995.  
XX  
XX 01-JUN-1995; 95CA-02150789.  
XX  
XX 02-JUN-1994; 94US-00252995.  
XX  
XX (MOUNT SINAI HOSPITAL CORP.  
XX  
XX Dennis JW, Heffernan M, Fode C;  
XX  
XX WPI; 1996-129817/14.  
XX P-PSDB; AAR92177.  
XX  
XX Nucleic acid encoding Sak serine-threonine kinase - useful for  
XX identifying modulators potentially useful in treatment or prevention of  
XX proliferative disease.  
XX  
XX Claim 5; Page 50-55; 73pp; English.  
XX  
XX A cDNA sequence (AAT08711) encodes sak-a (AAR92177), a novel  
XX serine/threonine kinase associated with mitotic and meiotic cell  
XX division. It was isolated from a murine lymphoid cell (D3JW25) cDNA  
XX library established in CHO cells by selection of wheatgerm agglutinin-  
XX resistant clones. Another cDNA clone (AAT08712) encoded the sak-b isoform  
XX (AAR9214). Sak-a and sak-b are probably alternatively spliced forms of  
XX the gene. The sak gene can be used for prodn. of recombinant SAK, as a  
XX probe in the diagnosis of proliferative disorders or, in antisense form,  
XX may be used to treat such disorders  
XX  
XX Sequence 3447 BP; 1061 A; 690 C; 754 G; 941 T; 0 U; 1 Other;

Query Match 64.5%; Score 1878; DB 2; Length 3447;  
Best Local Similarity 80.9%; Pred. No. 0;  
Matches 2361; Conservative 0; Mismatches 411; Indels 147; Gaps 9;  
QY 1 ATGGGCACTGATATGGGGAAGATCGAGATTTTAAAGTTGGAATCTGCTGTAA 60  
Db 206 ATGGCGGCGGTGATGGGGAAGATCGAGACTTTAAGTTGGAATCTGCTGTAA 265  
QY 61 GGATCATTTGCTGTCTACAGAGCTAGTCACTGCTTGAAGTTGCAATC 120  
Db 266 GGATCATTTGCTGTCTACAGAGCTAGTCACTGCTTGAAGTTGCAATC 325  
QY 121 AAAATGATGATGAAAGCCATGTACAAAGCAGAAATGTGACAGAGTCCAAATAG 180  
Db 326 AAAATGATGATGAAAGCCATGTACAAAGTGAATGTGACAGAGTCCAAATAG 385  
QY 181 GTGAAATATACATTCGCAATGAAACATCTCTATCTTGAGCTTTATACATTTGAA 240  
Db 386 GTGAAATATACATTCGCAATGAAACATCTCTCTGTGAGCTTTATACATTTGAA 445  
QY 241 GATAGCAATTAATGTGTATCTGTATTAAGAAATGTGCCATTAATGAGAAATGACAGTAT 300  
Db 446 GATAACAATTAATGTGTATCTGTATTAAGAAATGTGCCATTAATGAGAAATGACAGTAT 505  
QY 301 CTAAAGAAATAGAGAAACCTTCTCAGAAATGAAAGTGCACCTTCATGCAACGATC 360  
Db 506 CTGAAGAAAGAAATGAAAGCTTCTCAGAAAGGAAAGTGCACCTTCATGCAACGATC 565  
QY 361 ATCAGAGGATGTGTATCTTCATTCATGATGATGATACACCGGAGCTTCACATTTCT 420  
Db 566 ATCAGAGGATGTGTATCTTCATTCATGATGATGATACACCGGAGCTTCACATTTCT 625  
QY 421 AACCTCTACGATCTGTAATATGAATCAAGATTGCTGATTTGGGCTGGCACTCAA 480  
Db 626 AACCTCTACGATCTGTAATATGAATCAAGATTGCTGATTTGGGCTGGCAACGAG 685  
QY 481 CTGAAATGCGACATGAAAGCACTATATATGAAATCTCTATCTTCACTTCCACCA 540  
Db 686 TTGAATATGCGACATGAAAGCACTATATGAAATCTCTATCTTCACTTCCACCA 745  
QY 541 GAAATGCCACTGAGATGACATGGCTTGAATCTGATGTTGGTCCCTGGGCTGATG 600  
Db 746 GAAATGCCACTGAGATGACATGGCTTGAATCTGATGTTGGTCCCTGGGCTGATG 805  
QY 601 TTTATACATTAATTCGAGGAGACCACTTGGACCTGACCTGACCTGACCTGACCTG 660  
Db 806 TTTATACATTAATTCGAGGAGACCACTTGGACCTGACCTGACCTGACCTGACCTG 865  
QY 661 AATTAAGTATGATGAGATTAATGAATGCAATCTTTTGTCAATTAAGGCGCAAGAG 720  
Db 866 AATTAAGTATGATGAGATTAATGAATGCAATCTTTTGTCAATTAAGGCGCGAGAG 925  
QY 721 CTTATTCACAGTTACTTGTGTAATTCAGAGATGCTTGAATGCTGCTTCAATATG 780  
Db 926 CTTATTCACAGTTACTTGTGTAATTCAGAGATGCTTGAATGCTGCTTCAATATG 985  
QY 781 GACCAATCTTTTATGTCGCAATTCCTCAACAAAGTAAAGATTAGGAATGTGGAA 840  
Db 986 GACCAATCTTTTATGTCGCAATTCCTCAACAAAGTAAAGATTAGGAATGTGGAA 1045  
QY 841 GACTCAATGATGATGAGGAGACCAATTTCTACATTAACAGCTTCTTCAGTACC 900  
Db 1046 GACTCAATGATGATGAGGAGACCAATTTCTACATTAACAGCTTCTTCAGTACC 1105  
QY 901 AGATTAAGGTGATGTTTATTTGACAAAGAACTTTGATGTTGACCACTTCCAAAT 960  
Db 1106 AGTTTAAGGTGACCTTACTGAC--AGAAAGCTTTGTTGTTCAACACTTCCAAAT 1162  
QY 961 AAATGACTGATATTTCCAAAGATTAAGTCAACTGATTTTCTCTCAGAGATGGA 1020  
Db 1163 AAATGACTGATATTTCCAAAGATTAAGTCAAGTGACTT--TTCTTCAGAGATGGA 1219

1021 AACAGTTTATATCTAGTGGGAAT-----CAAGAAACAGTATAGTGAAGGGA 1074  
1220 AGTAATTTTGTATGATGGGAATCCAGAAACAGAAAGTATAGTGGGACGGGG 1279  
1075 AGAGTAATTCAGATGAGAGAAAGGCCAATTCGATACCTTGTAGAGTTATTC 1134  
1280 AAGAGTATGAGATGAGAGAGAGGCGCATTCGATACCTGCGAGAGCTATTC 1339  
1135 TGTGATGATCTGGCACTTCTAATAGTCTGATCAGAGCAAAACATATCAGTGA 1194  
1340 TGTGATGAGCAGCCCTCTTA--TCAGTCTGAGCAAAACATCAGTGAAGCT 1396  
1195 TGTCTCTGAGCAAAATGCTTCAAGTGTCAAAAGATCAGAGAGGTGAATGAAG 1254  
1397 TGTCTCTGAGTGAATGCTTCAAGCCTTGAAGATC----- 1434  
1255 AGTACTCACCAGACAGAACATGCCATTTTAACTTCTTAAAGAAAGATCC 1314  
1435 ----- 1434  
1315 AGTACTCTGATCTTTGAAAGCCTGATTAACATCAAGCACTCTCCATATCTTTGT 1374  
1435 -----ACTGATGAAATCAACAGTTTCAATCATCTTGT 1471  
1375 CCAGGAAAACTCTTTTCAATTTGAGAGCCGACCTCAGACTGAACCGTACAG 1434  
1472 CTAGGAAAACTCTTTTCAATTTGAGAGCCGACCTCAGTGAATGATGAGAG 1531  
1435 TGTGTTGGGAATCTGCAATTAATGCTCAATTAAGAAAACTAATGATGAGAG 1494  
1532 TGTGTTGGGAATCTGCAATTAATGCTCAATTAAGAAAACTAATGAGAGAG 1591  
1495 ACCCCAAACCGGAGCTTCCAGGCGCATCCAGATTTGAGAGAGAGCACTAA 1554  
1592 ACCCCAAAGAGATTTGAGAGCTATCCAGATTTGAGAGAGAGAGAGAGAG 1648  
1555 TGAAGTATGAGAGAGTCAAAAGTCAATGCTGATGATTAATGAGAGAGAG 1614  
1649 TGAAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1708  
1615 CAGCAAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1674  
1709 CAGCTGAG 1768  
1675 GATGATGTTTGGGCTCAGATCTCTTTGAGAGAGAGAGAGAGAGAGAGAG 1734  
1769 GA-----GCCGGGCTCATCTCTCATCTGAGAGAGAGAGAGAGAGAG 1822  
1735 CAGTGGGTTTATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1794  
1823 AAGTGGGTTTATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1882  
1795 TTAAG 1854  
1883 TTAAG 1942  
1855 GTGTGTGAG 1914  
1943 GTGTGTGAG 2002  
1915 ATATCTAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1974  
2003 ATATCTAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2062  
1975 CTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2034  
2063 CTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2122  
2035 CCAG 2094  
2123 CCAG 2182  
2095 AATCTCCAG 2154

2183 AAAAGTCCCAATATCACTATTTTACAGATATGCTAAATGATTTGATGAGAAATTC 2242  
2155 CCGTGTGCTGATTTTGAAGTTTGTATTAATGAGAGAGAGAGAGAGAGAGAG 2214  
2243 CCGTGTGCTGATTTTGAAGTTTGTATTAATGAGAGAGAGAGAGAGAGAGAG 2302  
2215 TTCAATCAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2274  
2303 TTCAATCAGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2362  
2275 AATAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2334  
2363 ACCAGCTGAAG 2422  
2335 TGTGATGAG 2394  
2423 TGTGATGAG 2482  
2395 TTCCCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2454  
2483 TTCCCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2542  
2455 CCGTCTTCTGTGATTTCAATTAACCAAGAGAGAGAGAGAGAGAGAGAG 2514  
2543 CCGTCTTCTGTGATTTCAATTAACCAAGAGAGAGAGAGAGAGAGAGAGAG 2599  
2515 ATGATGAG 2574  
2600 GTGAATGAG 2659  
2575 GAAG 2634  
2660 GAAG 2708  
2635 GATGATGCTTCTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2694  
2709 -----GCTTCTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2764  
2695 GCTACAG 2754  
2765 GCTACAG 2824  
2755 GTGAG 2814  
2825 GTGAG 2884  
2815 GAG 2874  
2885 GAG 2944  
2875 CTTGATGATTTTCTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2913  
2945 CTTGATGATTTTCTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2983

RESULT 9  
AAT08712  
ID AAT08712 standard; cDNA, 1599 BP.  
XX  
XX AAT08712;  
XX  
XX 25-MAY-1996 (first entry)  
XX  
XX DE Sak-b serine-threonine kinase cDNA.  
XX  
XX Sak-b; serine-threonine kinase; agonist; antagonist;  
XX KM proliferative disease; cancer; tumour; antisense; transgenic animal;  
XX therapy; se.  
XX OS  
XX Mub musculus.  
XX  
XX FH Key Location/Qualifiers

FT CDS 206..1599  
 FT /\*tag= a  
 FT /note= "T bases may also be U"  
 PN CA2150789-A.  
 XX  
 XX 03-DEC-1995.  
 PD  
 XX 01-JUN-1995; 95CA-02150789.  
 XX  
 XX 02-JUN-1994; 94US-00252995.  
 PR  
 PA (MOUN ) MOUNT SINAI HOSPITAL CORP.  
 PI Dennis JW, Heffernan M, Fode C;  
 XX WPI; 1996-129817/14.  
 DR P-PSDB; AAR92214.  
 XX  
 PT Nucleic acid encoding Sak serine-threonine kinase - useful for  
 PT identifying modulators potentially useful in treatment or prevention of  
 PT proliferative disease.  
 XX  
 PS Claim 6; Page 58-61; 73pp; English.

CC A cDNA sequence (AAT08712) encodes sak-b (AAR92214), a novel  
 CC serine/threonine kinase associated with mitotic and meiotic cell  
 CC division. It was isolated from a murine lymphoid cell (D3JW25) cDNA  
 CC library established in CHO cells by selection of wheatgerm agglutinin-  
 CC resistant clones. Another cDNA clone (AAT08711) encoded the sak-a isoform  
 CC (AAR92177). Sak-a and sak-b are probably alternatively spliced forms of  
 CC the gene. The sak gene can be used for prodn. of recombinant SAK, as a  
 CC probe in the diagnosis of proliferative disorders or, in antisense form,  
 CC may be used to treat such disorders

Sequence 1599 BP; 481 A; 342 C; 378 G; 398 T; 0 U; 0 Other;

Query Match 33.3%; Score 970.4; DB 2; Length 1599;  
 Best Local Similarity 85.0%; Pred. No. 3.5e-242;  
 Matches 1161; Conservative 0; Mismatches 186; Indels 19; Gaps 6;

QY 1 ATGGCGACCTGCATCGGGGAGAAATCGAGATTTTAAAGTTGAAATCGCTGGTAA 60  
 DB 206 ATGGCGGCGCTGCATCGGGGAGAGATCGAGATTTTAAAGTTGAAATCGCTGGTAA 265  
 QY 61 GGATCATTTGCTGGTGTCTTACAGAGCTGATCACTGCTGGTGGAGTTGCAATC 120  
 DB 266 GGATCATTTGCTGGTGTCTTACAGAGCTGATCACTGCTGGTGGAGTTGCAATC 325  
 QY 121 AAAATGATGATAGAAAGCCATGTACAAGAGAAATGCTACAGAGTCCAAATGAG 180  
 DB 326 AAAATGATGATAGAAAGCCATGTACAAGAGAAATGCTACAGAGTCCAAATGAG 385  
 QY 181 GTGAAAATGATATGTCCTTCTATCTTGAAGCTTTATTAATCTATTTTGA 240  
 DB 386 GTGAAAATGATATGTCCTTCTATCTTGAAGCTTTATTAATCTATTTTGA 445  
 QY 241 GATAGCATTTATGTGTATCTGTATTAAGAAATGTCCTTAATGAGAAATGAACAGTAT 300  
 DB 446 GATTAACAATTTATGTCTTACCTGTATGAAATGTGCCCAATGAGAAATGAACAGTAT 505  
 QY 301 CTAAAGATAGAGTGAACCTTCTCAGAAAATGAAAGCTGACCTTATGACCAAGATC 360  
 DB 506 CTGAAGAACAGATGAACCTTCTCAGAAAAGGAAAGCTTATGACCAAGATC 565  
 QY 361 ATCAGAGGATGTTGTATCTTCAATCTCATGTATTAACAACGAGCACTTCT 420  
 DB 566 ATCAGAGGATGTTGTATCTTCAATCTCATGTATTAACAACGAGCACTTCT 625  
 QY 421 AACGCTCAGTGAAGCTGTAATATGAACCAAGATGCTGATTTTGGCTGGCACTCAA 480  
 DB 626 AACATCTTACTTAAGCGGAAATGAACATAAAATTTGCTGACTTTGGAAGTCAAGCAG 685

QY 481 CTGAAAATGCGCATGAAAAGCATATACATTAATGTGAACCTCTTAACATCAATTCACCA 540  
 DB 686 TTGAATATGCGCACATGAAAAGCATATACATCTGTGGGACTCTAATTAATTCACCA 745  
 QY 541 GAAATGGCACTCGAAGTGCACATGAGCCCTTGAATCGTATGTTGGTCCCTGGCTGTATG 600  
 DB 746 GAAATGGCACTCGAAGTGCACATGAGCCCTTGAATCGTATGTTGGTCCCTGGCTGTATG 805  
 QY 601 TTTTATACATTAATTAATCGGAGACACCCCTTGCACATGACACAGTCAAGAACATTA 660  
 DB 806 TCTTATACCTTACTTAATGAAGACCACTTTTGAACATGACACAGTCAAGAACATTA 865  
 QY 661 AATTAAGTATATGCGAGATTAATGAATGCACTTTTGTCAATAGAGCCCAAGAC 720  
 DB 866 AACAAAGTATGCTGCGAGATTAATGAATGCAAGCTTTTGTCAAGAGCCCAAGAC 925  
 QY 721 CTTATTCACAGTTACTTCTGTAAGAAATCGACAGATCGTTTAAGCTCTTCAAGTATG 780  
 DB 926 CTTATTCACAGTTACTTCTGTAAGAAATCGACAGATCGTTTAAGCTCTTCTGTGTG 985  
 QY 781 GACCATCTTTTATGTCCGAAATCTTCAACAAAAGTAAAGATTAGAATCTGTGAA 840  
 DB 986 GACCATCTTTTATGTCCGAAATCTTCAACAAAAGTAAAGATGAGAGTGTAGAG 1045  
 QY 841 GACTCAATTTGATAGTGGGCGATGCCAAATTTCTAAGTGAATTAAGCTTTCCAGTACC 900  
 DB 1046 GACTCAATTTGATAGTGGGCGATGCCAAATTTCTAAGTGAATTAAGCTTTCTGTGATCC 1105  
 QY 901 AGTATAGTGTGATTTTATTTTGAACAAAAGACCTTTGATGTGTGACCACTCCCAAT 960  
 DB 1106 AGTTGATGTGCAACCTTCTGAC---AGAAAGCTTTTGTGTGCAACCTTCCAAAT 1162  
 QY 961 AAAATGACTGATTTTCAAAAAGAAATGAATGCTGATTTTCTTTCAGAGATGA 1020  
 DB 1163 AAAATGACTGATTTTCAAAAAGAAATGAATGCTGATTTTCTTTCAGAGATGA 1219  
 QY 1021 AACAGTTTATTAATCTAGTGGGAAAT-----CAAGAAACAGTAAATGGAAGGGA 1074  
 DB 1220 AGTAAATTTTGTATCTAATGGGAAATCCAGAACAAAGCTAATGATAGGGAAGGAGG 1279  
 QY 1075 AGATTAATTCAGATGCAAGAAAGGCAACATTTCTGATACCTTGTAGAGCTTATTC 1134  
 DB 1280 AGATTAATTCAGATGCAAGAAAGGCAACATTTCTGATACCTTGTAGAGCTTATTC 1339  
 QY 1135 TCTGATGATCTGSCACTTCTAATATGTCAGTCTCAAGCAAAAATATTAATGAAGCA 1194  
 DB 1340 TCTGATGATGAGCCAGCCCTCTTA---TCAATCTGAGCAAAAATATTAATGAAGCA 1396  
 QY 1195 TGTCACTGACGAAATGCTTTCAGTGTCCAAAAGATCAGAGAGGTGAATAATGAAG 1254  
 DB 1397 TGTCACTGACGAAATGCTTTCAGTGTCCAAAAGATCAGAGAGGTGAATAATGAAG 1453  
 QY 1255 AGGTACTACCCCAAGCAACATTTGCAATTTTAACTTTTAAAGAAAGACATCC 1314  
 DB 1454 AGGTACTACCCCAAGCAACATTTGCAATTTTAACTTTTAAAGAAAGACATCC 1512  
 QY 1315 AGTATCTGATCTTTTGAAGAAAGCTGATTAACATCAAGCACTCT 1360  
 DB 1513 AATAGTTAAGGATCTTTTGAAGAAAGCTGATTAACATCAAGCACTAT 1558

RESULT 10  
 AAT08710  
 ID AAT08710 standard; cDNA; 1453 BP.  
 XX  
 XX AAT08710;  
 AC  
 XX  
 XX  
 DT 25-MAY-1996 (first entry)  
 XX  
 XX Sak serine-threonine kinase N-terminus-encoding cDNA.  
 DE Sak; serine-threonine kinase; agonist; antagonist; proliferative disease;  
 KM cancer; tumour; antisense; transgenic animal; therapy; ss.

XX OS Mus musculus.  
 XX FH Key Location/Qualifiers  
 FT CDS 206..1453  
 FT /tag= a  
 FT /note= "T bases may also be U"  
 XX CA2150789-A.  
 XX 03-DEC-1995.  
 XX 01-JUN-1995; 95CA-02150789.  
 XX 02-JUN-1994; 94US-00252995.  
 XX (MOUN ) MOUNT SINAI HOSPITAL CORP.  
 XX Dennis JW, Heffernan M, Fode C;  
 XX WPI; 1996-129817/14.  
 XX P-PSDB; AAR92176.  
 XX Nucleic acid encoding Sak serine-threonine kinase - useful for  
 PT identifying modulators potentially useful in treatment or prevention of  
 PT proliferative disease.  
 XX Claim 4; Page 46-48; 73pp; English.  
 XX A cDNA sequence (AAT08710) encodes the N-terminal 416 amino acids  
 CC (AAR92176) of 2 isoforms, sak-a and sak-b, of a novel serine/threonine  
 CC kinase associated with mitotic and meiotic cell division; sak-a and sak-b  
 CC differ in their C-terminal sequences (see AAT08711 and AAT08712). cDNA  
 CC clones for sak-a and sak-b were isolated from a murine lymphoid cell  
 CC (D3W25) cDNA library established in CHO cells by selection of wheatgerm  
 CC agglutinin-resistant clones. Sak-a and sak-b are probably alternatively  
 CC spliced forms of the gene. The gene can be used for prodn. of recombinant  
 CC SAK, as a probe in the diagnosis of proliferative disorders or, in  
 CC antisense form, may be used to treat such disorders  
 XX SQ Sequence 1453 BP; 426 A; 315 C; 357 G; 355 T; 0 U; 0 Other;  
 Query Match 31.7%; Score 924.2; DB 2; Length 1453;  
 Best Local Similarity 86.4%; Pred. No. 3.7e-230;  
 Matches 1071; Conservative 0; Mismatches 153; Indels 15; Gaps 4;  
 QY 1 ATGGCAGCTGCATCGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGTAA 60  
 Db 206 ATGGCGCGCTGCATCGGGAGAGATCGAGATTTTAAAGTTGGAATCTGCTGTAA 265  
 QY 61 GGATCATTTTCTGTGTCTTACAGAGTGAATCATTCACATGCTTTGGAAGTTGCAATC 120  
 Db 266 GGATCATTTTCTGTGTCTTACAGAGTGAATCATTCACATGCTTTGGAAGTTGCAATC 325  
 QY 121 AAAATGATGATTAAGAAAGCATGTAAGAGAGAGATGTTGAGAGAGTCCAAAATGAG 180  
 Db 326 AAAATGATGATTAAGAAAGCATGTAAGAGAGATGTTGAGAGAGTCCAAAATGAG 385  
 QY 181 GTGAAATATCATTTGCAATTTGAAACATCTTCTATCTTGAGAGCTTTTAACTATTTTGA 240  
 Db 386 GTGAAATATCATTTGCAATTTGAAACATCTTCTGTTGAGAGCTTTTAACTATTTTGA 445  
 QY 241 GATAGCAATTTATGTTATCTGTATTAAGAAATGTCATTAAGAGAAATGAAAGAGTAT 300  
 Db 446 GATTAACAATTTATGTTATCTGTATTAAGAAATGTCATTAAGAGAAATGAAAGAGTAT 505  
 QY 301 CTAAAGAAATGAGTAAACCTCTCTGCAAAATGAAAGCTGCAACTTCATGACCAAGATC 360  
 Db 506 CTGAAGAAATGAGTAAACCTCTCTCTGCAAAATGAAAGCTGCAACTTCATGACCAAGAT 565  
 QY 361 ATCAGAGGATGTTGTTATCTTCAATTCATGATTAATCAACCGGAGCTCAACTTTCT 420  
 Db 566 ATCAGAGGATGTTGTTATCTTCAATTCATGATTAATCAACCGGAGCTCAACTTTCT 625

QY 421 AACCTCTACTGATCTGTAAATATGAAATCAAGATTTGCTGATTTTGGCTGGCACTCA 480  
 Db 626 AACATCTTACTTACCGGGAATATGAAATATTAATTTGCTGATTTTGGATTAACAGCAG 685  
 QY 481 CTGAAATATGCAATGAAAGCACTATATCATTTATGTAATCTTAACTAATTTTCA 540  
 Db 686 TTGAATATGCAATGAAAGCACTATATCATTTATGTAATCTTAACTAATTTTCA 745  
 QY 541 GAAATGCACTGCAAGTGCACATGAGCTTGAATCTGATGTTTGGTCTGGCTGTATG 600  
 Db 746 GAAATGCACTGCAAGTGCACATGAGCTTGAATCTGATGTTTGGTCTGGCTGTATG 805  
 QY 601 TTTTATACATTTCTATGCGGAGACCACTTGCACATGACAGTCAAGAACATTA 660  
 Db 806 TCTTATACATTTCTATGCGGAGACCACTTGCACATGACAGTCAAGAACATTA 865  
 QY 661 AATTAAGTATATGCGAGATTAATGAAATGCAATCTTTTGTCAATAGAGCCAAAGAC 720  
 Db 866 AACAAAGTATATGCGGAGATTAATGAAATGCAATCTTTTGTCAATAGAGCCAAAGAC 925  
 QY 721 CTATTCACAGTATCTTGTGAAATCCAGAGATGTTTAACTTGTCTTCAATAG 780  
 Db 926 CTATTCACAGTATCTTGTGAAATCCAGAGATGTTTAACTTGTCTTGTGTTG 985  
 QY 781 GACCATCTTTATATGCGGAAATCTTCAACAAAAGTAAAGATTAGAACTGTGAA 840  
 Db 986 GACCATCTTTATATGCGGAAATCTTCAACAAAAGTAAAGATTAGAACTGTGAA 1045  
 QY 841 GACTCAATGATATGCGGAGATCCAAATTTCTACTGCAATTAACGCTTCTTCAATAC 900  
 Db 1046 GACTCAATGATATGCGGAGATCCAAATTTCTACTGCAATTAACGCTTCTTCAATAC 1105  
 QY 901 AGTATAGTGTATGTTTATTTGACAAAGAACTTTGATGTTGAGGACCTCCCAAT 960  
 Db 1106 AGTATAGTGTATGTTTATTTGACAAAGAACTTTGATGTTGAGGACCTCCCAAT 1162  
 QY 961 AAAATGCTGATTTTCCAAAGATTAAGTTCAACTGATTTTCTTCTTCAAGAGATGA 1020  
 Db 1163 AAAATGCTGATTTTCCAAAGATTAAGTTCAACTGATTTTCTTCTTCAAGAGATGA 1219  
 QY 1021 AACAGTTTATATCTGATGCGGAAAT-----CAAGAAACAGTAAATAGTGAAGGGA 1074  
 Db 1220 AAGTAAATTTTGTATCTAATGCGGAAATCCAGAAAGAAAGTAAATAGTGAAGGGA 1279  
 QY 1075 AAGTAAATTTTGTATCTAATGCGGAAATCCAGAAAGAAAGTAAATAGTGAAGGGA 1134  
 Db 1280 AAGTAAATTTTGTATCTAATGCGGAAATCCAGAAAGAAAGTAAATAGTGAAGGGA 1339  
 QY 1135 TGTGATAGATCTGGCACTTCTAATAGTCTCAAGCAAAACATATCAATGAAAGCA 1194  
 Db 1340 TGTGATAGATCTGGCACTTCTAATAGTCTCAAGCAAAACATATCAATGAAAGCA 1396  
 QY 1195 TGTCACTAGCAAAATGCTTCAAGTGTCCAAAGATCA 1233  
 Db 1397 TGTCACTAGTGAATGCTTCAAGCTTAAGATCA 1335  
 RESULT 11  
 ACH34348  
 ID ACH34348 standard; cDNA; 484 BP.  
 XX ACH34348;  
 AC 13-OCT-2003 (first entry)  
 XX  
 DE Human endothelial cell cDNA #2481.  
 XX  
 KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
 XX genome mapping; biodiversity; genetic disorder.  
 OS Homo sapiens.  
 XX





QY 1608 TGTAAACAGCAAAATACATGATATATATGACCTTCAGACCTTAACCTGAGATAT 1667  
 DB 181 TGTAAACAGCAAAATACATGATATATATGACCTTCAGACCTTAACCTGAGATAT 240  
 QY 1668 CCAACAGAAATGTTTTGGCTCAGATCTCTTTCTGAACAGAGCAAGCTAGGGGTAT 1727  
 DB 241 CCAACAGAAATGTTTTGGCTCAGATCTCTTTCTGAACAGAGCAAGCTAGGGGTAT 300  
 QY 1728 GGAGCCACATGGGGTTATCATGAATGCTACA 1758  
 DB 301 GGAGCCACATGGGGTTATCATGAATGCTDNA 331

RESULT 13  
 ADD33867  
 ID ADD33867 standard; DNA; 609 BP.  
 AC ADD33867;  
 XX  
 XX 15-JAN-2004 (first entry)  
 DT  
 XX Mouse mitochondrial DNA sequence SEQ ID NO:1640.  
 DE  
 XX  
 XX de; mouse; array; mitochondrial; hybridisation; energy-metabolism;  
 KM mitochondrial disease; oxidative phosphorylation dysfunction;  
 KM oxidative stress; apoptosis; aging.  
 XX  
 OS Mus musculus.  
 XX  
 XX WO2003020220-A2.  
 PN  
 XX 13-MAR-2003.  
 PD  
 XX 30-AUG-2002; 2002WO-US027886.  
 PF  
 XX 30-AUG-2001; 2001US-0316323P.  
 PR 31-AUG-2001; 2001CA-02356540.  
 PR  
 XX (UYEM-) UNIV EMORY.  
 PA  
 XX Wallace DC, Levy S, Kerstann K, Procaccio V;  
 PI  
 XX MPI; 2003-300821/29.  
 DR  
 XX  
 XX Array containing probes for genes involved in mitochondrial biology,  
 PT useful for determining mitochondrial biology gene expression profiles for  
 PT use in diagnosing pathologies and identifying biochemical pathways.  
 PS  
 XX Claim 2; SEQ ID NO 1640; 201bp; English.

The invention relates to a novel array comprising at least two isolated nucleotide molecules, each molecule having a sequence capable of uniquely hybridizing to a nucleic acid molecule which is an expression product of a gene involved in mitochondrial biology. The array comprises two or more isolated nucleic acid molecules or spots, each molecule having a sequence chosen from sequence of 994 human probes and 2046 mouse probes. An array of the invention is useful for determining an expression profile of a mouse or human sample containing nucleic acid, by contacting the array with the sample under conditions allowing selective hybridisation, and measuring hybridisation of nucleic acid in the sample to the array to produce an expression profile. The array is also useful for determining an expression profile of a first labelled sample containing nucleic acid relative to a second, differently labelled sample containing nucleic acid. The second sample is a reference or a standard. An array is useful for determining an expression profile diagnostic of an energy-metabolism-related physiological condition. An array of the invention is useful for determining mitochondrial biology gene expression profiles of organisms, such as human, mice and closely related species, tissue and organs of such organisms, which are useful for determining expression profiles diagnostic of energy metabolism-related physiological conditions, diagnostic of such physiological conditions, identifying biochemical pathways, genes, and mutations involved in such physiological conditions, identifying therapeutic agents useful for preventing and/or treating such

CC physiological conditions, evaluating and/or monitoring the efficacy of  
 CC such therapies, and creating and identifying animal models of human  
 CC energy metabolism-related physiological conditions. An array is also  
 CC useful for defining expression signatures or profiles for mitochondrial  
 CC diseases, as well as distinguishing clinical disorders that result from  
 CC oxidative phosphorylation (OXPHOS) dysfunction, oxidative stress,  
 CC apoptosis and aging. An array of the invention contains probes of genes  
 CC not previously recognised to participate in mitochondrial biology. The  
 CC sequences shown in ADD3324-ADD35260 represent murine mitochondrial DNA  
 CC clones used to make the probes of the invention. Some sequences are not  
 CC present, these are SEQ ID NO's 295, 1174, 1213, 1700, 1728, 1730, 1905,  
 CC 1906, 2408 and 2643.  
 XX

Sequence 609 BP; 189 A; 121 C; 113 G; 186 T; 0 U; 0 Other;  
 SO

Query Match 9.7%; Score 284; DB 9; Length 609;  
 Best Local Similarity 82.1%; Pred. No. 1.8e-63;  
 Matches 345; Conservative 0; Mismatches 60; Indels 15; Gaps 1;

QY 2494 GCATCTTCAAGAAATGCTCAATGATGCTGCTTCTTCAACACAGGACCAATCCTT 2553  
 DB 6 GCCTCAGCAAGCAAGCTGAGCTGATATGTCGCTTCCCAACAGATCCCGAGACTC 65  
 QY 2554 AATCCCTCTATGCTTCAAAATGAAAGACCTGCTTCAACATACAGCTTCTGGAACAGAC 2613  
 DB 66 AGTCCTTCCACTGTGACAGTTGAAGACTTGGCAACAGCGACTGCCAAGGAACAGGC 125  
 QY 2614 ATCTCTTCAATATAGCTTAAAGATTTCTTCTTCAATACGACCACTTTGAATCTGTT 2673  
 DB 126 GTCTCTTCA-----GTCTCTTCAATCTGCAACAGCTTTTGAATCTGTT 170  
 QY 2674 TTGTGAAAAATGTTGGTGGCTACACAGTTAACTAGTGAAGCTGTGGTTCAGTTT 2733  
 DB 171 TTGTGAAAAATGTTGGTGGCTACACAGTTAACTAGTGAAGCTGTGGTTCAGTTT 230  
 QY 2734 AATGATGGTCCAGTTGTTGTGCAAGGAGAGTCTTCTTCACTAGTTATACCTGACCA 2793  
 DB 231 AATGATGGTCCAGTTGTTGTGCAAGGAGAGTCTTCTTCACTAGTTATACCTGACCA 290  
 QY 2794 AATGTCAAACAACTATGATATGAGAAATTAATCCAGACTACATCAACAGAA 2853  
 DB 291 GATGTCAGACAACTATGATATGAGAAATTAATCCAGACTACATCAACAGAA 350  
 QY 2854 TTACAGTGTCTGCTTCCATCTTCTTGAATTTCTTCACTCCGACTCTATTTTCATGA 2913  
 DB 351 TTACAGTGTCTTCTTCCATCTTCTTGAATTTCTTCACTCCGACTCTATTTTCATGA 410

RESULT 14  
 ABL12501  
 ID ABL12501 standard; cDNA; 2585 BP.  
 XX  
 XX ABL12501;  
 AC  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 31985.  
 DE  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX  
 XX Drosophila melanogaster.  
 OS  
 XX WO200171042-A2.  
 PN  
 XX 27-SEP-2001.  
 PD  
 XX 23-MAR-2001; 2001WO-US009231.  
 PF  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 PR  
 XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;  
 XX MPI; 2001-656860/75.  
 DR P-PSDB; ABB68398.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 XX Claim 1; SEQ ID NO 31985; 21bp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutic and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB150511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 SQ Sequence 2585 BP; 764 A; 596 C; 569 G; 656 T; 0 U; 0 Other;

Query Match 9.7%; Score 282.2; DB 4; Length 2585;  
 Best Local Similarity 60.0%; Pred. No. 1e-62;  
 Matches 470; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

QY 14 TCGGGGAGAGATCGAGATTTTAAAGTTGGAATCGCTTGTAAAGATCTTTGCTG 73  
 Db 104 TTGGAGAAACATTTAGAGACTTAAAGTACAGACTTGTGGTTAAAGGTTTTCGA 163  
 QY 74 GTGCTTACAGAGTCAAGTCCATTCACATGCTTGGAGTTGGAATCGAATCAATGATGATA 133  
 Db 164 CCGCTTCAAGAGGCGCGTGTCTGACATCAACAGATGCGCATTAAGATGATGATA 223  
 QY 134 AGAAGCCATGTACAAAGAGAGATGTACAGAGATCCAAATGAGTGAATACATT 193  
 Db 224 AAAAATCAATCCAGGCACTGAGCTACATCAACGCTTGGCCAGAGAGTGAATCCACT 283  
 QY 194 GCCAATTTGAAATCCTTCTATCTTGGAGCTTTATTAATTTTGAAGATAGCAATTATG 253  
 Db 284 CCGGCTGGAAGATCCCTCTGTGCTTCACTGTACACTTTCTTGAAGACGCAACTATG 343  
 QY 254 TGTATCTGATTTAGAAATGTGCTCAATATGAGAAATGAACAGTATCTTAAAGATAGAG 313  
 Db 344 TGTATTTGGTCTCGAGCTGCGCCATTAATGAGGAGCTTATGCTATATGAAACCAATCG 403  
 QY 314 TGAACCTTCTCAGAAATGAGCTCGACCTTCATGACACAGATCAACAGGAGTGT 373  
 Db 404 CCAGACCTTCAAGAGAGAGAGGCTCTTCAATTTGAGAGAGTGTGCGGAGACTTT 463  
 QY 374 TGTATCTTCAATCTCATGTATTAATCAACCGGAGCTCACTTTCTAACCTCTACTGA 433  
 Db 464 TGTACCTGACATCCCAACATCATGACACCGGAGATTTGCTGTCACCACTACTGCTTA 523  
 QY 434 CTCGTAAATGAAATCAAGATTTGCTGATTTTGGCTGGCAATCAATGAAATGCCAC 493  
 Db 524 GAGAGGAAATGACACTCAAGATGAGCCGATTTTGGCTGGCAATCAATGAAAGCACTG 583  
 QY 494 ATGAAAGCACTATATATGTGGAACCTCTTAATCATATTCAACAGAAATTCCTCACTC 553  
 Db 584 ATGACCGCATATATGACATATGTGGAATCCGAACTATATTTGCTGAGGTGTATGCG 643  
 QY 554 GAAAGTCACATGAGCTTGAATCTGATTTGCTGCTGAGCTGATATTTTATATCAATTAC 613  
 Db 644 GAACTCTTCAAGAGATGCGCGGAGCGTCTGAGAGGTGATGATGCTGATCAACCTGCG 703  
 QY 614 TTATTCGAGAGCAACCTTTCGACATGACATCAAGAACATTAATTAATTAATATATAT 673  
 Db 704 TGTGTGAGCGCGCGCTTGAACCGATGACATGACATCAACCTTAAACAAATGTGTGA 763

QY 674 TGGCAGATTATGAAATGCCATCTTTTGTCAATAGAGCGCAAGACCTTATTCACAGT 733  
 Db 764 TGTCCAGTACATTAATGCCGCTCACTTGTCTTACAGAGCGGAGACCTGATTAACAAAT 823  
 QY 734 TACTGTGAAATATCCAGAGATCGTTTAAGTCTGTCTTCAATTTGAGCAATCTTTTA 793  
 Db 824 TATTAAAAAGCTCCGACAGCAATTAACCTGAGAGCTGTCTCTGCCATCCCTTTA 883  
 QY 794 TGT 796  
 Db 884 TGT 886

RESULT 15

ACH37505  
 ID ACH37505 standard; cDNA; 491 BP.

ACH37505;

13-OCT-2003 (first entry)

Human endothelial cell cDNA #5638.

Human; ss; sequencing by hybridisation; SHH; expressed sequence tag; EST;

genome mapping; biodiversity; genetic disorder.

Homo sapiens.

US2003073623-A1.

17-APR-2003.

30-JUL-2001; 2001US-00918995.

30-JUL-2001; 2001US-00918995.

(DRMA/) DRMANAC R T.

(LABA/) LABAT I.

(STRAC/) STACHE-CRAIN B.

(DICK/) DICKSON M C.

(JONE/) JONES L W.

Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

WPI; 2003-615964/58.

Claim 1; SEQ ID NO 24717; 44bp; English.

The invention relates to an isolated polynucleotide comprising any one of  
 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
 determined by the technique of SBH (sequencing by hybridisation). Also  
 included is a purified polypeptide comprising a sequence corresponding to  
 a reading frame of the novel polynucleotide. The nucleic acid sequences  
 are useful in diagnostics as expressed sequence tags (EST) for  
 identifying expressed genes or for physical mapping of the human genome,  
 in forensics, in assessing biodiversity, or in identifying mutations  
 responsible for genetic disorders and other traits. The nucleotide  
 sequences are also useful as hybridisation probes, as oligomers for PCR,  
 for chromosome and gene mapping, in the recombinant production of  
 protein, or in generating antisense DNA or RNA. The purified polypeptide  
 is useful for generating antibodies specific for it. The present sequence  
 is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
 for this patent did not form part of the printed specification, but was  
 obtained in electronic format directly from USPTO at  
 seqdata.uspto.gov/sequence.html?docid=20030073623  
 Sequence 491 BP; 134 A; 104 C; 141 G; 106 T; 0 U; 6 Other;

Query Match 8.6%; Score 250; DB 8; Length 491;  
Best Local Similarity 100.0%; Pred. No. 1.2e-54;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGGGCACTGCATCGGGGAGAGATCGAGATTTTAAAGTTGAAATCTGCTGTAA	60
Db	241	ATGGGCACTGCATCGGGGAGAGATCGAGATTTTAAAGTTGAAATCTGCTGTAA	300
Qy	61	GGATCATTTGCTGTGCTCTACAGAGCTGATTCACACTGTGTTGGAAGTTGCAATC	120
Db	301	GGATCATTTGCTGTGCTCTACAGAGCTGATTCACACTGTGTTGGAAGTTGCAATC	360
Qy	121	AAATGATAGATAGAGAGCATGTACAAAGCAGAAATGTACAGAGATCCAAATGAG	180
Db	361	AAATGATAGATAGAGAGCATGTACAAAGCAGAAATGTACAGAGATCCAAATGAG	420
Qy	181	GTGAAAATACATTTGCCAATTGAAACATCCTTCTATCTTGAGCTTTATACTATTTGAA	240
Db	421	GTGAAAATACATTTGCCAATTGAAACATCCTTCTATCTTGAGCTTTATACTATTTGAA	480
Qy	241	GATAGCAATT	250
Db	481	GATAGCAATT	490

Search completed: September 24, 2004, 15:16:48  
Job time : 1097 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 14:00:23 ; Search time 7172 Seconds  
(without alignments)  
12128.909 Million cell updates/sec

Title: US-10-026-021-1  
Perfect score: 2913  
Sequence: 1 atcgagcactgcacgcgsgga.....cgactcccaatttcacatga 2913

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_escha:\*  
2: em\_eschum:\*  
3: em\_esctn:\*  
4: em\_esctmu:\*  
5: em\_esctov:\*  
6: em\_esctpl:\*  
7: em\_esctro:\*  
8: em\_htc:\*  
9: gb\_escl:\*  
10: gb\_escl2:\*  
11: gb\_htc:\*  
12: gb\_escl3:\*  
13: gb\_escl4:\*  
14: gb\_escl5:\*  
15: em\_esctun:\*  
16: em\_esctom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrt1:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1069.2	36.7	1656	11	AK006459 Mus muscu
2	925.8	31.8	3402	11	AK045082 Mus muscu
3	826.6	28.4	879	9	AU130274 AU130274
4	779.2	26.7	1182	11	AK006827 Mus muscu

5	750.4	25.8	825	14	CD558072 AGENCOURT
6	735.6	25.3	3137	11	BC045337 Dantio rer
7	685.4	23.5	937	12	BM903951 AGENCOURT
8	669.8	23.0	976	10	BF794900
9	623.2	21.4	1027	13	BF706811
10	620.6	21.3	860	9	AL705839
11	619.6	21.3	797	12	BG290478
12	616.4	21.2	691	12	BG255560
13	609	20.9	827	13	BU166217
14	571	19.6	816	13	BQ440730
15	567	19.5	878	14	CB196323
16	555	19.1	557	12	BM830522
17	554.2	19.0	870	10	BF215170
18	538.8	18.5	848	13	BU708651
19	537.8	18.5	874	10	BF245698
20	536.4	18.4	596	12	BM759840
21	534.6	18.4	585	9	AV716458
22	531.4	18.2	875	10	BF243227
23	526	18.1	770	12	BI829311
24	513.8	17.6	528	9	AL043645
25	512.4	17.6	660	14	CF914035
26	507.2	17.4	650	14	CA558867
27	500.4	17.2	643	14	CF915391
28	499	17.1	637	14	CF175127
29	499	17.1	638	14	CF914476
30	496.8	17.1	870	9	AI894043
31	494.8	17.0	559	9	AI435520
32	494	17.0	570	12	BMS37120
33	493.4	16.9	625	14	CF174552
34	491.8	16.9	506	9	AL048822
35	491.2	16.9	939	10	BF970082
36	489	16.8	619	14	CD552690
37	485.4	16.7	582	10	AW117220
38	479.6	16.5	787	14	CD643986
39	475.8	16.3	753	12	BG917180
40	471.6	16.2	592	14	CD553656
41	470.8	16.2	552	10	BE683420
42	466.8	16.0	590	14	CF906371
43	465.2	16.0	584	14	CA560592
44	463.6	15.9	584	14	CD561189
45	463.4	15.9	659	13	BY762320

#### ALIGNMENTS

RESULT 1  
AK006459  
LOCUS Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:11700028H20 product:serine/threonine kinase 18, full insert sequence.  
DEFINITION  
AK006459 1656 bp mRNA linear HTC 20-SEP-2003  
ACCESSION  
AK006459  
VERSION AK006459.1 GI:12839571  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE  
1  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
2  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20493374  
PUBMED 11042159

REFERENCE	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, X., Izawa, M., Ohgami, E., Matsubara, S., Kawai, J., Okazaki, Y., Ishikawa, T., Ozawa, X., Tanaka, N., Ohtani, T., Kashihara, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashi, Z.
JOURNAL MEDLINE PUBMED	RIKEN Integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
REFERENCE	20530913
PUBMED	11076861
AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
JOURNAL REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 1656)
JOURNAL REFERENCE	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukushima, Y., Furuno, M., Hanagaki, T., Hatanai, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Ota, C., Salto, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, Y., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan [E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216]
COMMENT	Please visit our web site ( <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> ) for further details.  cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and genome science laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAAGGAATCCCAAGATCGAAGCTCTTTTTTTTNN 3'], cDNA was prepared by using tritonalase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence[5'-GAGAGAAGAGGGCCGCCCATTAATTTCGTGAATTTAAATTAAATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI, 3' end: SstI. Host: SOLR.  location/Qualifiers 1..1656 /organism="Mus musculus" /mol_type="mRNA" /scriptain="C57BL/6J" /db_xref="FANTOM_DB:1700028H20" /db_xref="MGI:1910124" /db_xref="taxon:10090" /cldn="1700028H20" /sex="male" /tissue_type="testis" /clon_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult"

```

/note="unnamed protein product; putative
serine/threonine kinase 18 (MD)|MGI:101783, GB|AK006459,
evidence: BLASTn, 100%, match=1556"
/codon_start=3
/protein_id="BAB24599.1"
/db_xref="GI:12839572"
/translation="KTYSVSRCHSEVMKSPKRSLENDNHSNHHCLGKTPFPFADQTP
PQEMVQDMFGNLOMNAHLGEETNEHHIVSPNRQDPDLDITLRNMTPTIRASKND
TSAAVHAKOLSAKRYVSAHHKREVPVQRFQKPAHSEKNSRSMETLIGKQRTLLS
ITSLILAHRLKPIROKTKKAAVSLIDSEEVCEVLLRSCAEGYVKEVLQISDGETMT
VYVNDNGRGPFLADRPPLPTDNTISRYAFNDLPEKMYKQYASRFQIVLNSKPKITY
FTRAKCIIEMNSBGADFEVWFYDAGATHTENLIIHLEKGTSTYLNKNEAVTSLEK
EUYVMYHANEGRICLSLESVISEEKSRSGSPFTIVGRKRGNTSKPALVAPV
DPSCCKEQQASRLSNASAPLPQSPGLSPSTVYEGLTGHTATGTGCVSSIPKSA
QILKSFVFNKVGWATOLTSQAVVWQFDPGSLQLVVQAGVSSIVTSPDGQVTRVGENK
LPEYIKOKLOCLSSILMFPSNPNFQ"

```

**ORIGIN**

Query Match	36.7%	Score 1069.2	DB 11	Length 1656
Best Local Similarity	81.9%	Pred. No. 6,4e-246		
Matches 1291	Conservative	0	Mismatches 258	Indels 27
			Gaps	4

QY	1338	ACCTGATTAACAATCAAGCACTCTCCAATCATCTTGTGCCAGAAAAA	CTCTTCTTCATT	139	
Db	62	ACTGGATGAAAAATCAACACAGTCTCCAATCATCTGTCTAGAAAAA	ACTCTTTCATT	121	
QY	1398	TGCAGATCCCGACACCTCGACCTGAAAACCGTACACAGTGGTTTGG	GAATCTTCGAATATA	145	
Db	122	TGCAGACCGACACCTCGAGATGAAAATGGTACAGAGTGTTTGGAA	NTCTGCAATGAA	181	
QY	1458	TGCTCATTTTAAAAAAA	CTACTGAATATGACAGCATACAGCCCAACCGGGACCTTCAAGG	151	
Db	182	TGCTCATTTTAAAGAAACCTAATAGACACACACCGTTAGCCAAAC	AGATATTTCAAGA	241	
QY	1518	CCATCCAGATTTGGCGAAGAGACATCAAAAAATCCTGGA	CTGTATACAAAAGTCAAAA	1577	
Db	242	CTATCCAGATTTGGC---	AGACACGTTTACGAAACCGCTTGACTGACACAGAGCCAGCA	298	
QY	1578	GAACCTGATGGCTTGATTAATGACATTTCTGTAA	AACAGCAAAATACCATGAATATAT	1637	
Db	299	GAATCTGATTA	CTTTGCCAATGTTCAATGCTGTAAGACGCTGATGTCATGAATAATGT	358	
QY	1638	GACTGCACCTTCA	CACTGTAACCTGAGATATCCACAAAGATGTGTTTGGCTCGATCC	1697	
Db	359	GAGTGCACATCACCATTAAGCCTGAGGTATGCA	CAAGA-----GCCGGGCTACATCC	412	
QY	1698	TCTTTCTGAACAGACGAAC	CTAGGGGTATGAGACCACCATGGGGTTATACGAATCTGAC	1757	
Db	413	TCATTTCTGAACAAAGCAAGATTAAGATGAGATCGCA	CACTGGGTTTACGAAAACCTAC	472	
QY	1758	ATTAGAAGACTTATCATCTCCGTTGGTTGCTCA	CAGGTTAAACCATATCAGACGAAAAC	1817	
Db	473	CTTAAAGAGTTATTA	CAATCTCTCTGATGTCTCA	CAGATTTAAAGCCAATCAACAGAAAAC	532
QY	1818	CAAAAAGCGTGTGAGACAT	ACTTGTGATTCCAGAGAGGTGTGTGTGAGCTTGTAAAGGA	18777	
Db	533	CAAAAAGCGTGTGAGACAT	CTTGTATTCAGAGAGGTGTGTGTGAGCTTGTGAAGA	592	
QY	1878	GTATGCACTTCACAAATATGTGAAAAGAT	CTTCAAGATATCTAAGATGAGAAATACAT	1937	
Db	593	GTGTGCGCTCGAAGAGATATGTGAAAAGATGCTTCA	ATATTCAGATGTGGAGCTATGAT	652	
QY	1938	CACATATTTATATTC	CAATGGTGTGTAAGGTTTTCCCTTGCTGATATGACCACTTCAAC	1997	
Db	653	CACGTTTATTTACCCGAACGATGGAAGAGCTTTC	CTTGTGCTGACAGACCTCCCTTGCC	712	
QY	1998	TACTGACAAATCATCAGTATGATACAGCTTGTGA	CAATTTTACGAAAAAATATCGCGGAAATA	2057	
Db	713	TACTGACAAATCATCAGTATGATACAGCTTGTGA	CAATCTTACGAAAAAATATCGCGGGAATA	772	
QY	2058	TCATATATGCTTCCAGGTTTGTACAGCTTGTAAATCT	TAAATCTCCAAAATCACTATTT	2117	
Db	773	TCAGTATGCTTCCAGATTCATTCAGCTAGTAAATCT	TAAATCTCCAAAATCACTATTT	832	



QY	2118	TACAAAGATATGCTAAATGCACTTTTGTGATGAGAAATCTCCTGGTGTGATTTTGAGCTTG	2177
Db	833	TACAAAGATATGCTAAATGATATTTTGATGGAATAATCTCTGGTGTGATTTGAAAGTTTG	892
QY	2178	GTTTATGATGGGTAAAAATACACAAACAGAAATTTCAATTCAGTGATTTGAAAGAC	2237
Db	893	GTTTATGATGAGAGCAAAATACATTAATCTGAAAAATTTAATTCATTAATTGAGAAAC	952
QY	2238	AGGGAAGCTTACACTTTAAAAAGTGAAGGAAATTAATAGCTTGAAAGAGAGATTA	2297
Db	953	AGGGAATACCTTATATTTAAAAAAGAAATGAAGATTAACAGCCTGAAAGAGGAATGAA	1012
QY	2298	AATGATATGAGACCATGCTAATGAGGGTCATGATTTGTTTGAACATGGAATCCATAT	2357
Db	1013	AGTATATATGAGCAATGCTAATGAGGGTCACCGATTTGCTTGTCATCTGGAATCTGTAT	1072
QY	2358	TTTCAGAAAGAGAAAGAAAACTAGGAGTGCTCCCTTTTCCCAATTAATCATAGAGAA	2417
Db	1073	CTCTGAGAGAGAAAGAAAGAACAGAGGGGTCTTCAATTCCTCCCTAATATCTAGAGAA	1132
QY	2418	ACCTGATATGATCTAGTTTCACTTAAGCCTTAATGACCTCCTCTCTGTGGAATCAAAATTA	2477
Db	1133	ACCTGATATATCTAGTTTCACTTAAGCCTTAATGACCTCCTCCTGGAACCAAGCTG	1189
QY	2478	CCCAACGAGAGATATAGAGCATTTTCAACAGATGGTCATGCAATAGTGTGCTTCTCCAAC	2537
Db	1190	CTGTATAGGAGAGAGCAGCGCTCAGACAGACAGACTGAGCGTGAATGTGCGCTTCCCAAC	1249
QY	2538	ACAGGACCAATTCCTTAATCCCTCTATGCTTACAAATGAGAGACTTGGTCTTTCAACTAC	2597
Db	1250	ACAGTGTCCCAAGACACTGCTCCCTTCCACTGTGTACAGTTGAAAGACTGTGCCACACAGCCAC	1309
QY	2598	AGCTTCTGGAAACAGACATCTCTTCAATATAGTCTAAAAAGATTTGTCTTCTTAATACAGACA	2657
Db	1310	TGCCACAGAAACAGCGCTCTCTTCAA-----GTCTTCTTAATCTGCACA	1354
QY	2658	ACTTTTGAATCTGTTTTTGTGAATAAATGTTGGTGGCTACACAGTTAACTAGTGAGC	2717
Db	1355	GCTTTTGAATCTGTTTTTGTGAATAAATGTTGGTGGCTACACAGCTAACTAGGGAGC	1414
QY	2718	TGTGTGGGTTCAGTTTAATGATGGGTCCCACTTGGTTGTGCAGGACAGAGTGTCTTCAAT	2777
Db	1415	TGTGTGGGTTCAGTTTAATGATGGGTCCCACTTGGTTGTGCAGGACAGAGTATCTTCCAT	1474
QY	2778	CAGTTATATCTCATCAAAATGCTCAAAACAATAGTATGAGAAAAATGAAAAATTAACAGA	2837
Db	1475	CAGTTATCAATCAACCAAGTGTCAACAACATAGTATGAGAAAAATGAAAAATTAACCTGA	1534
QY	2838	CTACATCAAAACAGAAATTAACAGTGTCTGTCTTCACATCCCTTTGATGTTTCTTAATCCGAC	2897
Db	1535	ATACATCAAAACAGAAATTAACAGTGTCTTCTTCACATCCCTCTGATGTTTCTTAATCCAAC	1594
QY	2898	TCCCTAATTTTCATGCA 2913	
Db	1595	TCCCTAATTTTCAGTGA 1610	

RESULT 2	
AK045082	
LOCUS	AK045082
DEFINITION	Mus musculus 9.5 days embryo parthenogenote cDNA, RIKEN full-length enriched library, clone: B130030L15 product:serine/threonine kinase 18, full insert sequence.
ACCESSION	AK045082
VERSION	AK045082.1 GI:26090702
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (house mouse)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 Carninci, P. and Hayashizaki, Y.
AUTHORS	

TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Mech. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	
AUTHORS	2
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE	Genome Res. 10 (10), 1617-1650 (2000)
PUBMED	20499374
REFERENCE	11042159
AUTHORS	3
	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

TITLE	JOURNAL MEDLINE PUBMED	REFERENCE AUTHORS
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Yamawake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, U., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	RIKEN integrated sequence analysis (RISA) system—384-format sequencing pipeline with 384 multicapillary sequencer	Genome Res. 10 (11), 1757-1771 (2000)
	20530913	
	11076861	
		4
		The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
		Functional annotation of a full-length mouse cDNA collection
		Nature 409, 685-690 (2001)

**TITLE** Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
**JOURNAL** Nature 420, 563-573 (2002)  
**REFERENCE** 6 (bases 1 to 3402)  
**AUTHORS** Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Horii, P., Imcwtari, K., Ishii, Y., Itoh, M., Kageawa, I., Kasukawa, T.,  
Kotchi, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koyas, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakanamura, M., Nishi, K., Nomura, K., Numasaki, R., Onno, M., Ohnato, N.,  
Okazaki, I., Salto, R., Saitoh, H., Sakai, K., Sakai, K., Sakaizume, N.,  
Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, K., Takahashi, F., Takahashi, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yaeunishi, A.,  
Miyamatsu, M., and Hayashizaki, Y.

**JOURNAL**  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, url: http://genome.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)  
**COMMENT**  
cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of genome Exploration Research Group in Riken Genomic Science Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Dr. Tomohiro Kono (Department of Animal Science, Tokyo University of Agriculture, 1737 Hnasei Atsugi City, Kanagawa, Prefecture, Japan) whose assistance we gratefully acknowledge.

```
FEATURES
  source
    Location/Qualifiers
      1. 3402
        /organism="Mus musculus"
        /mol_type="mRNA"
```

```

/strain="C57BL/6J"
/db_xref="PANTOM.DB:B130030L15"
/db_xref="MGI:2410441"
/db_xref="taxon:10090"
/clone="B130030L15"
/issue_type="parthenogenote"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="9.5 days embryo"
1..3402
misc_feature
/note="serine/threonine kinase 18 (MD|MGI:101783,
GB|AK06459, evidence: BLASTN, 100%, match=1656)"

```

```

Query Match
Best Local Similarity 31.8%; Score 925.8; DB 11; Length 3402;
Matches 1072; Conservative 0; Mismatches 152; Indels 15; Gaps 4;

```

```

QY 1 ATGGGAGCTGATCGGGAGAGAGATCGAGATTTTAAAGTTGAAATCTGCTTGTAA 60
DB 228 ATGGCGGCTGATCGGGAGAGAGATCGAGATTTTAAAGTTGAAATCTGCTTGTAA 287
QY 61 GATCATTTTGTGTGTCTCAAGAGCTGAGTCAATCACTGGTTTGAAGTTGCAATC 120
DB 288 GATCATTTTGTGTGTCTCAAGAGCTGAGTCAATCACTGGTTTGAAGTTGCAATC 347
QY 121 AAAATGATGATAGAAAGGCAATGTAAGAGAGAAATGATGAGAGAGTCCAAATGAG 180
DB 348 AAAATGATGATAGAAAGGCAATGTAAGAGAGAAATGATGAGAGAGTCCAAATGAG 407
QY 181 GTGAAATATATGATGCAATGCAATGCAATCTTCTATCTTGAAGCTTTTAACTATTGAA 240
DB 408 GTGAAATATATGATGCAATGCAATGCAATCTTCTATCTTGAAGCTTTTAACTATTGAA 467
QY 241 GATAGCAATATATGATCTGTGATTTGAAGATGTCATTAATGAGAAATGAGAGATAT 300
DB 468 GATAGCAATATATGATCTGTGATTTGAAGATGTCATTAATGAGAAATGAGAGATAT 527
QY 301 CTAAAGATATGATGAAAGCCCTCTCAAGAAATGAGAGCTGACACTTCTATGACCAATC 360
DB 528 CTAAAGATATGATGAAAGCCCTCTCTCAAGAAAGAGAGAGCTGACACTTCTATGACCAATC 587
QY 361 ATACAGGATGATGATCTTCTATCTTCAATGATATCAACCGGAGCTGACACTTCT 420
DB 588 ATACAGGATGATGATCTTCTATCTTCAATGATATCAACCGGAGCTGACACTTCT 647
QY 421 AACCTCTACTGCTGATATATGAAATCAATGATGATGATGATGATGATGATGATGATGAT 480
DB 648 AACCTCTACTGCTGATATATGAAATCAATGATGATGATGATGATGATGATGATGATGAT 707
QY 481 CTGAAGATGCAATGAAAGCACTATATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 708 CTGAAGATGCAATGAAAGCACTATATGATGATGATGATGATGATGATGATGATGATGAT 767
QY 541 GAAATGCACTGAGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 768 GAAATGCACTGAGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 827
QY 601 TTTTATACATTAATGATGAGAGCACTGATGATGATGATGATGATGATGATGATGATGAT 660
DB 828 TTTTATACATTAATGATGAGAGCACTGATGATGATGATGATGATGATGATGATGATGAT 887
QY 661 AATAAGATGATGAGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 888 AATAAGATGATGAGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 947
QY 721 CTATATCCAGATTAATGATGAGAGCACTGATGATGATGATGATGATGATGATGATGATGAT 780
DB 948 CTATATCCAGATTAATGATGAGAGCACTGATGATGATGATGATGATGATGATGATGATGAT 1007
QY 781 GACCATCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 1008 GACCATCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1067

```

```

QY 841 GACTCAATGATAGTGGGATGCGACAAATTTCTACTGCAATTAACAGCTTCCAGTACC 900
DB 1068 GACTCAATGATAGTGGGATGCGACAAATTTCTACTGCAATTAACAGCTTCCAGTACC 1127
QY 901 AGATTAAGTGTATGTTTATTTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 1128 AGTTAAGTGTATGTTTATTTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1184
QY 961 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1185 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1241
QY 1021 AACAGTTTATTAATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074
DB 1242 AGTAAATTTTATTAATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1301
QY 1075 AGAGTAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1134
DB 1302 AGAGTAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1361
QY 1135 TCTGATAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1194
DB 1362 TCTGATAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1418
QY 1195 TGTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1233
DB 1419 TGTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1457

```

```

RESULT 3
AUI30274 879 bp mRNA linear EST 01-AUG-2002
LOCUS AUI30274 NT2RP3 Homo sapiens cDNA clone NT2RP300533 5', mRNA
DEFINITION
sequence.
ACCESSION
AUI30274
KEYWORDS
AUI30274.1 GI:10990628
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

```

```

REFERENCE
1 (Bases 1 to 879)
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
Ishigai, T.
HRI human cDNA project
JOURNAL
Unpublished (2000)
COMMENT
Contact: Takao Iwogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

```

```

HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Location/Qualifiers
1..879

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP300533"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RP3"
/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"

```

```

Query Match
Best Local Similarity 28.4%; Score 826.6; DB 9; Length 879;
Matches 98.5%; Pred. No. 1.2e-187;

```







Qy	429	ACTGACTGTATATGAACTCAAGATTGCTGATTTTGGGCTGGCACTCACTGAAT	488
Db	440	GCTCACCAACGAGATGAACTTTAAATGCGAGCTTTGGTTTACCACTCCAACTCAAGCT	499
Qy	489	GCCACATGAAAAAGCATATACATTATGTGAACTCTTAATCACTTTCACCAAGAAATGCG	548
Db	500	TCCACAGTGAAGAGCATTTTACATGTGGGCAACCCAACTACATTTTCACTCCGAAGTGGC	559
Qy	549	CACCTGAAGTGCATGAGCCCTGTAATCTGANTGTTGGTCCCTGGGCTGTATGTTTATAC	608
Db	560	CACCGGAGGCGCTACCGGCTGTAGCTCAGACGTGTGCTCACTGGGCTGACATGTTTACGC	619
Qy	609	ATTACTATGGGAGACCACTCTTCACTGACACAGTCACAGTCACAACTTTAAATTAAGT	668
Db	620	CTTTCTCAGAGAGACCGGCTTTTGAACAGACCGGTCAACGCACTCCATTAAGT	679
Qy	669	AGTATGGGAGATTAATGAAATGCAATCTTTTGTCTCATATGAGGCCAAGACCTTATCA	728
Db	680	GGTGTGGAGAGATTAATGAAATGCCCATGACATATCAGAGAGGCTCAGAGATCTAATCCA	739
Qy	729	CCAGTACTTCTGTAATCCAGAGATCGTTTAAAGTCTGTCTTCAGTATTTGAGACCATCC	788
Db	740	GCAAGCTGTACAAAGAACCCAGCCCTCAGACCGAGCTTTCTGCAAGTGTGAGCAACCC	799
Qy	789	TTTTATGTCCCGAATTTCTTCAACAAAAGTAAAGATTTAGAACTGTGGA--AGACTC	845
Db	800	GTTTCATGACTCAAGTGGGCCCAACCGACAAAGACTCCGGCAGCAGACAGGAGCTC	859
Qy	846	AATTGATAGGGGATGCGCAATTTCTACTGCAATTAAGCTTCTTCCAGTACAGTAT	905
Db	860	CATCGACAGGGAGATGCTACATATCACAGC--TTCAATGCAACCAACCAACGACAG	916
Qy	906	AACTGTGATTTATTTGACAAAAGAACTTTTGTATGTGTGAGCCACTCCCAATTAAT	965
Db	917	CACAGTGTGGTTACAGAGGGGAAACAGGAGATGATTTGGGCAACCGTTCCTCAGTGTAT	976
Qy	966	GACTGATTTCCAAAGATTAAGTTCACTGATTTTCTTCTTCAAGAGATGAGAAACAG	1025
Db	977	GGACCAATCTTAGCCACTCCCATATTTCCACAGCTCAAGTTTAAAGTGGGAGAGA	1036
Qy	1026	TTTTTAATCTAGTGGGAAATCAAGAAACCAATTAAGTGAAGGGGAAAGTAAATCA	1085
Db	1037	CTGGGAGCCAAATTCACAGATGATTTATCCAGAGTGGGCGTTGGAGAAATCCCGTGGG	1096
Qy	1086	AAGTGAAGAAAGAGCCCACTTTCTGATATCTTGTGAGCTTATCTCTGATATGATC	1145
Db	1097	TGCTGACAGCGGAGGCTCATCTCTGCTAACGAAGAGCTCATTTGCTGATTCGATC	1156
Qy	1146	TGGCACTTCTAATATGCTCAGTCTCAGCAAAACATATACATGGAACGATGTCACTCAGC	1205
Db	1157	TGCTGTGGGTACAGCCACATATCCACAGAGGCAAG--GCTGGAGAGATGTCACTCAGA	1213
Qy	1206	AAGAAATGCTTTCAGTGTCCAAAGATCAGAGAGAGGTGAAATGAAGAGAGTCACTCACC	1265
Db	1214	AAGAGTGTCTTTCAGGGGCTGACGGTATTTTCCACAGACCTCCGCTATGAATATGCAACC	1273
Qy	1266	CACAGACAACATGCGCAACTTTTAACTTTTAAAGAAAAGATCCAGTATGTTGCG	1335
Db	1274	ACATGG-----AACTCCAAACATGACAGG	1288
Qy	1326	ATCTTTTGAAGACCTGATTAACATCAAGCACTCTCAATCATCTTTTCCAGGAAAAAC	1385
Db	1289	CTTCCCTCTCCACCTGTAAACCACTGTATCTCGGCAATGTCTTTCTTACGT---C	1354
Qy	1386	TCCTTTTTCATTTGACAGCCGACACTCAGACTGAAAACGATACACAGTGTGTTGGAA	1445
Db	1355	CACCTATTTCAAGAGACAACAAATGCCGATTTCTCAAC--ACAGCCCTGTTCAGGAA	1411
Qy	1446	TCTGCAATTAATGCTCATTTAAGAAAAAATGATGAAATATATACAGCATCAGCCAAACCG	1505
Db	1412	T-----GATGAGTGTTCAGAGGCGGCGAGACATGAGCGGTCAACAGCAGAGTGG	1462

Qy	1506	GGACTTCCAGGGCATCCAGATTTTGCAAGAGACATCAAAAAATGCTGTGACTGATAC	1565
Db	1463	GAACTTTCACCTGTGACGAG-----GCCAATCGGAGACAAACCTCTGACGTGAGCA	1516
Qy	1566	AAAAGTCAAAAAGAACTCTGAGTCTTCTGATTAATGCACTTCTGTAAAACAGCAAAATAC	1625
Db	1517	GCCCAAGGCTCTTCACTTCA--GCAGCAAGCCATCTCTTCCAGACAAACACCCGGGCG	1575
Qy	1626	CATGAATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAAAGAAATGTGTTT	1685
Db	1576	CCTGCAAGGAGATGATCTTGTATCAGGACACATGTCAAGACTCA-----GGCTTAT	1628
Qy	1686	TGGCTCAGATCTCTTCTTCTGAAACAGACAACTAGGGGTATGAGCCACATGGGTTA	1745
Db	1629	TCGACGCAAGTGTTCATGTTCACCGCTTCCAAAGGCAAGCAACAGAGAAAG	1688
Qy	1746	TCGAATGATCATTAAGAAAGCATTAATCACTCCGTTGGTGTGCTCAGGTTAAACCAAT	1805
Db	1689	GACAAAGTGTCTTAAGAAAGCTTC--CTTCGCTCTGTCTGCGACACTCAAGCTTAT	1747
Qy	1806	CAGACAGAAAACAAAAAGCTGTGTGAGCAATCTTGATTCAGAGAGGTGTGTGGA	1865
Db	1748	TCGACAGAAAGCTAAAAATGCTGTGTGAGCAATTTGGGCAATGAGAGGTGTATGGA	1807
Qy	1866	GCTGTAAAGAGATGATGATCTCAAGAAATATGTGAAGAAATCTTGAGATATCTATGGA	1925
Db	1808	GCTGTGAAAGAGACAAAGGGGCTCAGAGAGAGTCAAGAAATCTTGAGATATCTGTGA	1867
Qy	1926	TGAAATATGATCATATTTATTAATCCAAATGTGTGATGAGTTCCTCTGTGCTATAG	1985
Db	1868	TGATCTATGTGTGACAGTGTACAGCTTAATGAAGAAAGGGTTCCTGTGTGATCA	1927
Qy	1986	ACCACTTCACTCTGACAAACATCAGTATACGCTTTGACAAATTTACCAAAAAATA	2045
Db	1928	TCACCTTCAACCTTAAGACATCTCATCTGAGCTTTGATGATTTTACAGAAAAATA	1987
Qy	2046	CTGGCAAAATATCAATATGCTTCAGGTTTGTACACTGTGAATCTAATATCTCCAA	2105
Db	1988	CTGGAAAGATATCATGATACGCAACAAAGTTTGTGACACTGTCAATCTAAGACTCMAA	2047
Qy	2106	AATCATTTATTTCAAGATATGCTAAATGATTTTGTATGAGAAATCTCTGAGTGTGA	2165
Db	2048	AGTCACTTTATACCAAAATTTGCAAGTCAATGTGATGAGAACTCTCTAATCAGA	2107
Qy	2166	TTTTGAGTGTGTTTATGATGGGTTAAATATCAAAAAAGAGATTTCAATCTCAGT	2225
Db	2108	CTGGAAGTGTGCTTCTATATGTGTGCAAAACCATTAAGACAAAGTGAAGAGTGGCGT	2167
Qy	2226	GATTAAGAAAGACAGGAGGCTTACACTTTTAAAGTGAAGTGAAGTAAATAGCTTGA	2285
Db	2168	GGTGAAGAAAGTGGGAAGTCAATACAGTGAAGGGTTACGTTGGTGTGAGTGGGCTCA	2227
Qy	2286	AGAGGAATTAATAATGATATGACATGCTAATGAGGTCATGATTTTGTATGACT	2345
Db	2228	TCTGAATGCACTCATATTAAGCTTTCTGAGGAGGACATCGAATGTGTCTT	2287
Qy	2346	GGAATCCATATTTTCA---GAAGGAAAGGAAAACTAGAGTGTCTCCCTTTTCCCAAT	2402
Db	2288	GGAAGCTGCAATCAGCTGTGAGAGCAGCGCAGGCTTAAACATCTCATTTCTTCCAT	2347
Qy	2403	AATCATGGAAGAAACCTGTGTATGATGTTTCACTTAAAGGCTTATACCTTCTCTTC	2462
Db	2348	CACCATGTGGAGAGGCTG-----TGAACCGGAGTCTCTCAAGCTCC	2389
Qy	2463	TGTGATTTCAATTTTACCAACGAGATGAGAGATTTTCAACAGAAAGGTATGATG	2522
Db	2390	CGCTCTTCAATTTCAATCTTCTGCGCTCAGGGGTGACAGAAAGTGTCAAGTGTG	2449
Qy	2523	TGCTGCTTTCACACAGGACCAATCTTAAATCTCTATGTTTAAATGAAGAACT	2582
Db	2450	TCTCAAGTCTCGGACAGATCTCAATC---ACTCATCAATGATTTCTATGCGGGGTC	2506
Qy	2583	TGGCTTAACATCAAGCTTCTGGAACAGACATCTTCTTAATAGTTAAAGATGTCT	2642

Db 2507 AGACTTGACACAGCCAGGTGGCTTAAGCAGCTCCCGGTCCACAAAGATGAACGAC 2566  
 Qy 2643 TCTTAATACAGCAACTTTTGAATCTGTTTGTGAAAAATGTTGGTCTACCA 2702  
 Db 2567 CGTTAATCTGGAAAGTCTCAAGTCCATTTTGTGCGCAACATGATGGGCTCTCA 2626  
 Qy 2703 GTTAATAGTGAAGCTGTGGGTGAGTTAATGATGGGTCCAGTTGGTGTGAGAGC 2762  
 Db 2627 GCTAACCATGGGAGAGTCTGGGTGAGTTAATGATGGGTCCAGTGTGATGAAGT 2686  
 Qy 2763 AGAGTGTCTTATCATGTTATACCTCACCMAATGTCACAACTAGATGAGAGAA 2822  
 Db 2687 TGGCGTCTTGGATCATTTTATCTCCCGAGAGACATCACAAGGATTAAGAGAA 2746  
 Qy 2823 TGAATAATACAGACTACATCAACAAATTAAGTGTCTGTCTTCCATCTTTGAT 2882  
 Db 2747 CGAAGATTAACAGACTGATGAGAAAGAACTCACTGTGTCCAGATCTTGGGTCT 2806  
 Qy 2883 GTTTTCTATCC 2894  
 Db 2807 TCTGCCAATCC 2818

RESULT 7  
 BM903951 937 bp mRNA linear EST 12-MAR-2002  
 LOCUS AGENCOUNT 6695921 NIH\_MGC\_67 Homo sapiens cDNA clone IMAGE:5492458  
 DEFINITION 5', mRNA sequence.

ACCESSION BM903951  
 VERSION BM903951.1 GI:19353918  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 937)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM12114 row: 1 column: 11  
 High quality sequence stop: 637.

FEATURES  
 source Location/Qualifiers

1. 937  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5492458"  
 /isue\_type="telomblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1b="NIH MGC 67"  
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.75 kb. Library constructed by Life  
 Technologies."

ORIGIN

Query Match 23.5%; Score 685.4; DB 12; Length 937;  
 Best Local Similarity 95.7%; Pred. No. 1,1e-153;  
 Matches 738; Conservative 0; Mismatches 26; Indels 7; Gaps 3;  
 Qy 1544 CAAAAATCCTGAGTATACAAAGTCAAAAGAACTCTGATGCTTGTGATATGAC 1603  
 Db 24 CAAAAATCCTGAGTATACAAAGTCAAAAGAACTCTGATGCTTGTGATATGAC 83

Qy 1604 ATTCTGTAAGACGCAAAATATCAATGAATATATGACTGCACTTCACTAAACCTGAGA 1663  
 Db 84 ATTCTGTAAGACGCAAAATATCAATGAATATATGACTGCACTTCACTAAACCTGAGA 143  
 Qy 1664 TAATCCAAAGAAATGTTTTTGGTCTGATCTCTTTCTGAACAGACAAAGTAAAG 1723  
 Db 144 TATCCAAAGAAATGTTTTTGGTCTGATCTCTTTCTGAACAGACAAAGTAAAG 203  
 Qy 1724 GATGAGGCCACATGGGTTATCAGATCTGATCAATTAAGAACATTAATCTCCCTGG 1783  
 Db 204 GATGAGGCCACATGGGTTATCAGATCTGATCAATTAAGAACATTAATCTCCCTGG 263  
 Qy 1784 TTGCTCAGAGTTAAACCAATCAAGACAAACCAAAAGCTGTGTGAGACTACTTG 1843  
 Db 264 TTGCTCAGAGTTAAACCAATCAAGACAAACCAAAAGCTGTGTGAGACTACTTG 323  
 Qy 1844 ATTCAGAGAGGTGTGTGTGAGCTTTGAAGAGTATGATCTCAAGATATGTGAAG 1903  
 Db 324 ATTCAGAGAGGTGTGTGTGAGCTTTGAAGAGTATGATCTCAAGATATGTGAAG 383  
 Qy 1904 AAGTTCTTCAATATCTGATGATGAGAAATACGATCTATTAATCAATATGTGTA 1963  
 Db 384 AAGTTCTTCAATATCTGATGATGAGAAATACGATCTATTAATCAATATGTGTA 442  
 Qy 1964 GAGTTTCTCTGCTGATGAGACCACTCACTACTGACAAACATCAGTATGAGTCA 2023  
 Db 443 GAGTTTCTCTGCTGATGAGACCACTCACTACTGACAAACATCAGTATGAGTCA 502  
 Qy 2024 TTGACAAATTTACAGAAAAATATCTGCGAAATATCAATATGCTTCAGGTTTGA 2083  
 Db 503 TTGACAAATTTACAGAAAAATATCTGCGAAATATCAATATGCTTCAGGTTTGA 562  
 Qy 2084 TTGTAAGATCTAATCTCCCAATATCACTTATTAACAAATATGTAATGATTTGA 2143  
 Db 563 TTGTAAGATCTAATCTCCCAATATCACTTATTAACAAATATGTAATGATTTGA 622  
 Qy 2144 TGAGAAATCTCTGCTGATGATGATTTTGAAGTTTGTGTTATGATGGGTAAATAC 2203  
 Db 623 TGAGAAATCTCTGCTGATGATGATTTTGAAGTTTGTGTTATGATGGGTAAATAC 682  
 Qy 2204 AAACGAAATTTCAATTCAGTATGATGATGATGATGATGATGATGATGATGAT 2260  
 Db 683 AAACGAAATTTCAATTCAGTATGATGATGATGATGATGATGATGATGATGAT 742  
 Qy 2261 GTGAAGTAAATGATTAATGATGATGATGATGATGATGATGATGATGATGAT 2308  
 Db 743 GTGAAGTAAATGATTAATGATGATGATGATGATGATGATGATGATGATGAT 793

RESULT 8  
 BF794900 976 bp mRNA linear EST 12-JAN-2001  
 LOCUS 602256767P1 NIH\_MGC\_85 Homo sapiens cDNA clone IMAGE:4339985 5',  
 DEFINITION mRNA sequence.

ACCESSION BF794900  
 VERSION BF794900.1 GI:12099954  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 976)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov  
 Tissue Procurement: Louis Staudt, M.D., Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be



RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res.



QY 61 GGATCATTTGCTGTGTCTACAGAGCTGATGTCATTCACTGCTTTGGAAGTTGCAATC 120  
| | | | |  
Db 267 GGATCATTTGCTGTGTCTACAGAGCTGATGTCATTCACTGCTTTGGAAGTTGCAATC 326  
QY 121 AAAATGATGATGAGAAAGCCATGTCAGAAAGAGAGATGTCAGAGAGTCCAAATGAG 180  
| | | | |  
Db 327 AAAATGATGATGAGAAAGCCATGTCAGAAAGAGAGATGTCAGAGAGTCCAAATGAG 386  
QY 181 GTGAAATATGATGTCAGAAAGCCATGTCAGAAAGAGATGTCAGAGAGTCCAAATGAG 240  
| | | | |  
Db 387 GTGAAATATGATGTCAGAAAGCCATGTCAGAAAGAGATGTCAGAGAGTCCAAATGAG 446  
QY 241 GATAGCAATTAATGCTATCTGTATTTAGAAATGTCATTAATGAGAAATGAGAAAGATG 300  
| | | | |  
Db 447 GATAGCAATTAATGCTATCTGTATTTAGAAATGTCATTAATGAGAAATGAGAAAGATG 506  
QY 301 CTAAAGGATGATGAGAAAGCCCTTCTCAGAAATGAGCTGACATTCATGACAGATC 360  
| | | | |  
Db 507 CTAAAGGATGATGAGAAAGCCCTTCTCAGAAATGAGCTGACATTCATGACAGATC 566  
QY 361 ATCAAGAGATGCTGTATCTTCTCATTCATGATGATTAATGACAGAGCTGACATTCCT 420  
| | | | |  
Db 567 ATCAAGAGATGCTGTATCTTCTCATTCATGATGATTAATGACAGAGCTGACATTCCT 626  
QY 421 AACTCTCTACTGCTGTATTAATGACATGACATGCTGATTTGGGCTGCAACTCAA 480  
| | | | |  
Db 627 AACTCTCTACTGCTGTATTAATGACATGACATGCTGATTTGGGCTGCAACTCAA 686  
QY 481 CTGAAATGCACTGAAAGGACCTATATGATTAATGAGAACTCTTAATGATTCACCA 540  
| | | | |  
Db 687 CTGAAATGCACTGAAAGGACCTATATGATTAATGAGAACTCTTAATGATTCACCA 746  
QY 541 GAAATGCTCACTGAAAGGACCTATATGATTAATGAGAACTCTTAATGATTCACCA 599  
| | | | |  
Db 747 GAAATGCTCACTGAAAGGACCTATATGATTAATGAGAACTCTTAATGATTCACCA 806  
QY 600 GTTTTATGATTAATGATTAATGAGAACTCTTAATGATTCACCA 649  
| | | | |  
Db 807 GTTTTATGATTAATGATTAATGAGAACTCTTAATGATTCACCA 857

RESULT 11  
BG290478 797 bp mRNA linear EST 21-FEB-2001  
LOCUS BG290478  
DEFINITION 60238837F1 NIH\_MGC\_93 Homo sapiens cDNA clone IMAGE:4517096 5',  
mRNA sequence.  
ACCESSION BG290478  
VERSION BG290478.1 GI:13047365  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 797)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-rcmail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://imgc.lnl.gov  
plate: LNL0409 row: j column: 09  
High quality sequence stop: 649.  
Location/Qualifiers

FEATURES  
SOURCE

1..797  
/organism="Homo sapiens"  
/mol\_type="mRNA"

/db\_xref="taxon:9606"  
/clone="IMAGE:4517096"  
/issue\_type="transitional cell papilloma, cell line"  
/lab\_host="DHIOB (phage-resistant)"  
/clone\_id="NIH\_MGC\_93"  
/note="Organ: bladder; Vector: pCMV-Sport6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 Kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

Query Match 21.3%; Score 619.6; DB 12; Length 797;  
Best Local Similarity 94.4%; Pred. No. 6.8e-138;  
Matches 697; Conservative 0; Mismatches 34; Indels 7; Gaps 5;

QY 2093 CTAAATCTCCCAAAATCACTTATTTTACAGATATGCTTAAATGCACTTTTGAATGAGAAAT 2152  
| | | | |  
Db 1 CTAAATCTCCCAAAATCACTTATTTTACAGATATGCTTAAATGCACTTTTGAATGAGAAAT 59  
QY 2153 CTCTGTGCTGATTTTGAAGTTTGTATTAATGAGAGGCTTAAATACAAAGAGAG 2212  
| | | | |  
Db 60 CTCTGTGCTGATTTTGAAGTTTGTATTAATGAGAGGCTTAAATACAAAGAGAG 118  
QY 2213 ATTTGATTCAGGTGATTTGAAAGAGAGAGAGGCTTCACTTTTAAAGAGAGAGAG 2272  
| | | | |  
Db 119 ATTTGATTCAGGTGATTTGAAAGAGAGAGAGGCTTCACTTTTAAAGAGAGAGAG 178  
QY 2273 TTAATAGCTTGAAGAGAGATTAATAATGATTAATGAGAGAGAGGCTTCACTTGA 2332  
| | | | |  
Db 179 TTAATAGCTTGAAGAGAGATTAATAATGATTAATGAGAGAGGCTTCACTTGA 238  
QY 2333 TTTGTTTGAAGAGAGAGATTAATAATGATTAATGAGAGAGGCTTCACTTGA 2392  
| | | | |  
Db 239 TTTGTTTGAAGAGAGAGATTAATAATGATTAATGAGAGAGGCTTCACTTGA 298  
QY 2393 TTTTCCCAATGATTAATGAGAGAGAGAGGCTTGAATTAATGAGAGAGGCTTCACT 2452  
| | | | |  
Db 299 TTTTCCCAATGATTAATGAGAGAGAGAGGCTTGAATTAATGAGAGAGGCTTCACT 358  
QY 2453 CTCTCTCTCTGATGATTAATGAGAGAGAGAGGCTTGAATTAATGAGAGAGGCTTCACT 2512  
| | | | |  
Db 359 CTCTCTCTCTGATGATTAATGAGAGAGAGAGGCTTGAATTAATGAGAGAGGCTTCACT 418  
QY 2513 TCATGATGATGCTGCTTCTCCCAAGAGAGAGAGAGGCTTGAATTAATGAGAGAGGCTTCACT 2572  
| | | | |  
Db 419 TCATGATGATGCTGCTTCTCCCAAGAGAGAGAGAGGCTTGAATTAATGAGAGAGGCTTCACT 478  
QY 2573 ATGAAGAGCTTGTGCTTCACTACAGAGCTTCTGAAGAGAGAGGCTTGAATTAATGAGCTTA 2632  
| | | | |  
Db 479 ATGAAGAGCTTGTGCTTCACTACAGAGCTTCTGAAGAGAGAGGCTTGAATTAATGAGCTTA 538  
QY 2633 AAGATTTGCTTCTTAAATGAGAGAGAGGCTTGAATTAATGAGAGAGGCTTGAATTAATGAGCTTA 2692  
| | | | |  
Db 539 AAGATTTGCTTCTTAAATGAGAGAGAGGCTTGAATTAATGAGAGAGGCTTGAATTAATGAGCTTA 597  
QY 2693 GGGCTACAGATTAATGAGAGAGGCTTGAATTAATGAGAGAGGCTTGAATTAATGAGCTTA 2752  
| | | | |  
Db 598 GGGCTACAGATTAATGAGAGAGGCTTGAATTAATGAGAGAGGCTTGAATTAATGAGCTTA 657  
QY 2753 TTGTG--CAGGAGAGAGGCTTCTTCACTGATTAATGAGAGAGGCTTGAATTAATGAGCTTA 2808  
| | | | |  
Db 658 GGTGTTGAGAGAGAGAGGCTTCTTCACTGATTAATGAGAGAGGCTTGAATTAATGAGCTTA 717  
QY 2809 AGGTATGAGAGAGAGAGGCTTCTTCACTGATTAATGAGAGAGGCTTGAATTAATGAGCTTA 2826  
| | | | |  
Db 718 AGGTATGAGAGAGAGAGGCTTCTTCACTGATTAATGAGAGAGGCTTGAATTAATGAGCTTA 735

RESULT 12  
BG255560 691 bp mRNA linear EST 13-FEB-2001  
LOCUS BG255560  
DEFINITION 602367854F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4476061 5',

mRNA sequence.  
 accession BG255560  
 version BG255560.1 GI:12765376  
 keywords EST.  
 source Homo sapiens (human)  
 organism Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 691)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-r@mail.nih.gov  
 Tissue Procurement: DCTD/DTF  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLM10302 row: 1 column: 14  
 High quality sequence stop: 680.  
 Location/Qualifiers  
 1..691  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4476061"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_1lb="NIH\_MGC\_91"  
 /note="Organ: prostate; Vector: pCMV-Sport6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.4 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

ORIGIN  
 Query Match 21.2%; Score 616.4; DB 12; Length 691;  
 Best Local Similarity 98.4%; Pred. No. 3.9e-137;  
 Matches 665; Conservative 0; Mismatches 6; Indels 5; Gaps 4;

QY 935 TTTTATTTGGTCCAGGCACTCCCAATTAATGCTGATTTCCAAAGATTAAGTTCAA 994  
 |||||  
 DB 1 TTTAGATTGGTCAGCCACTCCCAATTAATGCTGATTTCCAAAGATTAAGTTCAA 60  
 |||||

QY 995 CGATATTTCTCTTCAGAGATGAAACAGTTTATCTCAGTGGGAAATCAAGAAA 1054  
 |||||  
 DB 61 CTGATTTTCTCTTCAGAGATGAAACAGTTTATCTCAGTGGGAAATCAAGAAA 120  
 |||||

QY 1055 CCAATATATGAGGAGGAGGAGATTAATCAAGATGCAAGAAAGCCCATTTTCGAT 1114  
 |||||  
 DB 121 CCAATATATGAGGAGGAGGAGATTAATCAAGATGCAAGAAAGCCCATTTTCGAT 180  
 |||||

QY 1115 ACCTTGTGAGGTTATTTCTCTGATGATGCTGCACTTCTATATGCTCAGTCCAGCAA 1174  
 |||||  
 DB 181 ACCTTGTGAGGTTATTTCTCTGATGATGCTGCACTTCTATATGCTCAGTCCAGCAA 240  
 |||||

QY 1175 AAACATATCAATGAAAGATGCTCAGTCCAGCAAGAAATGCTTCAAGTGTCCAAAGATCAG 1234  
 |||||  
 DB 241 AAACATATCAATGAAAGATGCTCAGTCCAGCAAGAAATGCTTCAAGTGTCCAAAGATCAG 300  
 |||||

QY 1235 GAGGAGGTGAAATGAGAGAGTACTCAACCAAGCAACAATGCAATTTTAACT 1294  
 |||||  
 DB 301 GAGGAGGTGAAATGAGAGAGTACTCAACCAAGCAACAATGCAATTTTAACT 360  
 |||||

QY 1295 TCTTTAA-AGAAAGACATCAGTACTGATCTTTTGAAGAGCTGATTAACATCA 1353  
 |||||  
 DB 361 TCTTTAAAGAAAGACATCAGTACTGATCTTTTGAAGAGCTGATTAACATCA 420  
 |||||

QY 1354 GCACCTCCCAATCATCTTTGTCCAGGAAATCTCTTT-TCCATTTGAGAGCCGACACC 1412  
 |||||

DB 421 GCACCTCCCAATCATCTTTGTCCAGGAAATCTCTTTCCATTTGACAGCCGACACC 480  
 |||||

QY 1413 TCAGACTGAACCGTACCAAGTGTGGGAAATCGCAATTAATGCTATTAAGAAA 1472  
 |||||

DB 481 TCAGACTGAACCGTACCAAGTGTGGGAAATCGCAATTAATGCTATTAAGAAA 540  
 |||||

QY 1473 AACTACTGAATATGACAGATGAGCCCAACCGGAGCTTCCAGGAGCTCAGATTTGCA 1532  
 |||||

DB 541 AACTACTGAATATGACAGATGAGCCCAACCGGAGCTTCCAGGAGCTCAGATTTGCA 600  
 |||||

QY 1533 GAGGACACATCAAA-AAATGCTTGCAGTATGCAAAA-GTCAAAAGACTCTGATGC 1589  
 |||||

DB 601 GAGGACACATCAAA-AAATGCTTGCAGTATGCAAAAAGTCAAAAGACTCTGATGC 660  
 |||||

QY 1590 TTCTGATATGACAT 1605  
 |||||

DB 661 TTCTGATATGACAT 676  
 |||||

## RESULT 13

BU166217 817 bp mRNA linear EST 04-SRP-2002  
 LOCUS AGENCOURT\_7952535 NIH\_MGC\_55 Homo sapiens cDNA clone IMAGE:6107893  
 5', mRNA sequence.

accession BU166217  
 version BU166217.1 GI:22680169

keywords EST.  
 source Homo sapiens (human)

organism Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 817)

NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Clontech Laboratories, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLM2351 row: e column: 14

High quality sequence stop: 562.

Location/Qualifiers

1..817  
 /organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6107893"

/tissue\_type="from acute myelogenous leukemia"

/lab\_host="DH10B (T1 phage-resistant)"

/clone\_1lb="NIH\_MGC\_55"

/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);  
 Site 1: SfiI (ggcgccctggcc); Site 2: SfiI

(ggccatctggcc); Double-stranded cDNA was prepared from

cell line RNA. 5' and 3' adaptors were used in cloning as

follows: 5' adaptor sequence: 5'-CACGCCATTTATGGCC-3' and

3' adaptor sequence:

5'-ATTCTAGAGCCGAGCGCCGACCATG-dT(30)BN-3' (where B = A,

C, or G and N = A, C, G, or T) Average insert size

1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained

inserts by PCR. This library was enriched for full-length

clones and was constructed by Clontech Laboratories (Palo

Alto, CA.)

ORIGIN  
 Query Match 20.9%; Score 609; DB 13; Length 817;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-135;  
 Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	2305	ATGACCACTGCTAATGAAGGGTCATCGTAATTGTTTAGCACTGGAAATCCATAATTTCCAGA	2366
Db	3	ATGACCANTCTATATGAGGGTCACTGCATTTGTTTGACACTGGAAATCCATAATTTCCAGA	62
Qy	2365	GAGGAAAAGAAAACTAGAGAGTGCCTCTTTTCCCAATATCATATGGAAGAAAACCTGGT	2422
Db	63	GAGGAAAAGAAAACTAGAGATGCTCCTCTTTTCCCAATATCATATGGAAGAAAACCTGGT	122
Qy	2425	AGTACTAAGTTCACTTAAGGCTTTATCATACCTCTCTCTTGAGATTCAAATTAACCAACG	2484
Db	123	AGTACTAAGTTCACTTAAGGCTTTATCATACCTCTCTCTTGAGATTCAAATTAACCAACG	182
Qy	2485	AGAGATAGAGCATCTTTCAACAGAAATGATCAGCAATAGTCTCTTCTTCCACA CAGGCA	2544
Db	183	AGAGATAGAGCATCTTTCAACAGAAATGATCAGCAATAGTCTCTTCTTCCACA CAGGCA	242
Qy	2545	CCAATCCTTAATCCCTCTATGTTACAAATGAAAGACTTGCTTACAACTACAGCTTCT	2604
Db	243	CCAATCCTTAATCCCTCTATGTTACAAATGAAAGACTTGCTTACAACTACAGCTTCT	302
Qy	2605	GGAACAGACATCCTCTCTAATAGTCTTAAAGATTTGTTCTTCTTAATCAGCAACTTTTG	2664
Db	303	GGAACAGACATCCTCTCTAATAGTCTTAAAGATTTGTTCTTCTTAATCAGCAACTTTTG	362
Qy	2665	AAATCTGTTTTGTGMAAATGTGGTGGGCTCACAGTTAACATGAGAGCTGTGG	2724
Db	363	AAATCTGTTTTGTGMAAATGTGGTGGGCTCACAGTTAACATGAGAGCTGTGG	422
Qy	2725	GTTTCAGTTTAATGATGGGTGCCAGTTGGTTGTGACGAGAGAGTCTTCTATCAGTTAT	2784
Db	423	GTTTCAGTTTAATGATGGGTGCCAGTTGGTTGTGACGAGAGAGTCTTCTATCAGTTAT	482
Qy	2785	ACCTGACCCAAATGCTCAACAACATCTAGTATGAGAAATGAAAAATTACACGACTAATC	2844
Db	483	ACCTGACCCAAATGCTCAACAACATCTAGTATGAGAAATGAAAAATTACACGACTAATC	542
Qy	2845	AAACAGAAATTACAGTGTCTGTCTTCCATCCTTTGATTTTCTAAATCCGACTCTAAT	2904
Db	543	AAACAGAAATTACAGTGTCTGTCTTCCATCCTTTGATTTTCTAAATCCGACTCTAAT	602
Qy	2905	TTTCATTGA 2913	
Db	603	TTTCATTGA 611	
<hr/>			
RESULT 14			
BQ440730			
LOCUS			
DEFINITION AGENCOURT 7890535 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6168199			
5', mRNA sequence.			
BQ440730			
VERSION BQ440730.1 GI:21179806			
KEYWORDS EST.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
NIH-MGC http://mgc.nci.nih.gov/.			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
Contact: Robert Strausberg, Ph.D.			
Email: cgabbs-rc@mail.nih.gov			
Tissue Procurement: ATCC/DCTP			
CDNA Library Preparation: Life Technologies, Inc.			
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Agencourt Bioscience Corporation			
clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LHAM3531 row: f column: 08			
High quality sequence stop: 745.			
Location/Qualifiers			

```

source
1. .826
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6168199"
/lissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-Sport6; site_1: NotI;
site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match      19.6%; Score 571; DB 13; Length 826;
Best Local Similarity 100.0%; Pred.No. 3.3e-126;
Matches 571; Conservative 0; Mismatch 3; Indels 0; Gaps 0

QY      1 ATGGCAGCTGTCATCGGGGAGAAGATCGAGATTAAAGTTGAAGTCTGTGGTAAA 60
DB      254 ATGGCCACCTGCATCGGGAGANATCGAGATTTAAGTTGAATCTGCTGTGAAA 313
QY      61 GGATCAATTTGCTGTGTCTTACAAGCTGATCATTCACACTGTTTGGAAGTTGCATC 120
DB      314 GGATCAATTTGCTGTGTCTTACAAGCTGATCATTCACACTGTTTGGAAGTTGCATC 373
QY      121 AAAATATATGATAAAGAACCATATCAAAGCAGAAATGTCACAGAGTCCAAATGAG 180
DB      374 AAATATATGATAAAGAACCATATCAAAGCAGAAATGTCACAGAGTCCAAATGAG 433
QY      181 GTGAAATATCATTTGCCAATTTGAAACATCTCTTATCTTGAGACTTATPACTATTTGAA 240
DB      434 GTGAAATATCATTTGCCAATTTGAAACATCTCTTATCTTGAGACTTATPACTATTTGAA 493
QY      241 GATAGCAATTATGTGTATCTGGTATTTAGAATGTGCCATATGAGAAATGAACAGTAT 300
DB      494 GATAGCAATTATGTGTATCTGGTATTTAGAATGTGCCATATGAGAAATGAACAGTAT 553
QY      301 CTAAAGATAGAGTGAAGAACCCCTTCCAGAAATGAAGCTCCGACATTGATGACCAAGATC 360
DB      554 CTAAAGATAGAGTGAAGAACCCCTTCCAGAAATGAAGCTCCGACATTGATGACCAAGATC 613
QY      361 ATCACAGGAGATTGTATCTTCATTTCTCATGGTATATACACCAGGACTTCACACTTCT 420
DB      614 ATCACAGGAGATTGTATCTTCATTTCTCATGGTATATACACCAGGACTTCACACTTCT 673
QY      421 AACCTCTACTGACCTCCGTAATATGAAATCAAAGTGTGTGATTTGGGCTGGCAACTCA 480
DB      674 AACCTCTACTGACCTCCGTAATATGAAATCAAAGTGTGTGATTTGGGCTGGCAACTCA 733
QY      481 CTGAAATATGCCACATGAAGAACCATATCATTTATGTGGAACTCCTAATCAATTTACCA 540
DB      734 CTGAAATATGCCACATGAAGAACCATATCATTTATGTGGAACTCCTAATCAATTTACCA 793
QY      541 GAATTTGCCATCTCGAATGTCACATGGCCCTTG 571
DB      794 GAATTTGCCATCTCGAATGTCACATGGCCCTTG 824

RESULT 15
CB196323
LOCUS      CB196323               878 bp    mRNA          linear   EST 05-FEB-2003
DEFINITION AGNCOCURT.11262941 NIH MGC 135 Mus musculus cDNA clone
IMAGJ:30138118 5', mRNA sequence.
ACCESSION  CB196323
VERSION     CB196323.1 GI:28223890
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 878)
            NIH-MGC http://www.ncbi.nlm.nih.gov/

```

TITLE  
JOURNAL  
COMMENT  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Dr. David Rowe  
CDNA Library Preparation: Invitrogen Corp  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILNL at:  
http://image.llnl.gov  
Plate: NDA0040 row: 9 column: 23  
High quality sequence spot: 676.  
Location/Qualifiers

## FEATURES

Source

1. 878  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30138118"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NH MGC 135"  
/note="Vector: pCMVSPORT6.1, Site 1: EcoRV, Site 2: NotI;  
Normalized full-length enriched library from pooled mouse  
embryonic limb, maxilla and mandible, day 12.5, 13.5,  
14.5, and 15.5 (size selected for the 0.5-1 kb fragments)  
Cloned directionally, priming method: Oligo-dT. cDNA  
enrichment: >1k bp. Average insert size 1.6k bp.  
Normalization (Cot value): 7.5 kb. Priming sequence:  
5'-GACTAGTTCGATGCGGAGCGGCCCTTT3' Tissue contributed  
by, David Rowe. Library constructed by Reagen, Invitrogen  
Corp."

## ORIGIN

Query Match 19.5%; Score 567; DB 14; Length 878;  
Best Local Similarity 82.3%; Pred. No. 3e-125;  
Matches 664; Conservative 0; Mismatches 140; Indels 3; Gaps 1;  
QY 1809 ACAGAAACCAAAAGGCTGTGTGAGCATCTGATTCAGAGAGAGGTGTGAGGCT 1868  
DB 1 ACAGAAACCAAAAGGCTGTGTGAGCATCTGATTCAGAGAGAGGTGTGAGGCT 60  
QY 1869 TGTAAAGAGATGATCTCAAGAAATATGTGAAGAAGTTCTTCAGATATCTAGATGG 1928  
DB 61 TGTGAAGAGGTGTGAGTGAAGATATGTGAAGAAGTCTTCAGATATCTAGATGG 120  
QY 1929 AAATACGATCACTTATTTATTCAGAAATGTGTGAGAGTTTCTCTTGTGATGACC 1988  
DB 121 GACTATGATCACTGTTATTAACCCGAAACGATGAAAGGCTTCTCTGTGACAGACC 180  
QY 1989 ACCCTCACTACTGACAAATAGATGATGATGATGATGATGATGATGATGATGATG 2048  
DB 181 TCCCTGCTACTGACAAATAGATGATGATGATGATGATGATGATGATGATGATG 240  
QY 2049 GCGAAATATCATATGCTTCAGAGTTTGTACAGCTTGTAGATCTAAATCTCCAAAT 2108  
DB 241 GCGGAAATATCATATGCTTCAGATTCATTCAGCTAGTAAATCTCCAAAT 300  
QY 2109 CACTTATTTTACAAGATATGCTAAATGATGATGATGATGATGATGATGATGATG 2168  
DB 301 CACTTATTTTACAAGATATGCTAAATGATGATGATGATGATGATGATGATGATG 360  
QY 2169 TGAAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2228  
DB 361 CGAAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420  
QY 2229 TGAAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2288  
DB 421 TGAAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480  
QY 2289 GGAATATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2348  
DB 481 GGAATATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540

QY 2349 ATCCATTAATTTGAG 2408  
DB 541 ATCTGTAATCTCTGAG 600  
QY 2409 AGGAG 2468  
DB 601 AGGAG 657  
QY 2469 TTCAATTATCCAG 2528  
DB 658 CCAAGCTCTCTTAAG 717  
QY 2529 TTCTCCACACAG 2588  
DB 718 TTCTCCACACAG 777  
QY 2589 TACAACTACAGCTTCTGGAACGACAT 2615  
DB 778 CACAGGAGCTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 804

Search completed: September 24, 2004, 20:22:01  
Job time : 7184 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model1

Run on: September 24, 2004, 14:58:38 ; Search time 1305 Seconds  
(without alignments)  
11305.898 Million cell updates/sec

Title: US-10-026-021-1

Perfect score: 2913  
Sequence: 1 atcgagcactgcacgcgggga.....cgacctaatcttcattga 2913

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2913	100.0	2913	15	US-10-026-021-1
2	2911.4	99.9	3937	15	US-10-037-270-280
3	2911.4	99.9	3937	16	US-10-117-722-280
4	2908.2	99.8	3721	13	US-10-116-802-108
5	2786.4	95.7	2836	13	US-10-425-114-26254
6	2686.8	92.2	3924	16	US-10-062-674-1868
7	426	14.6	484	10	US-09-918-995-21560
8	328	11.3	328	9	US-09-878-178-96
9	328	11.3	328	14	US-10-046-935-96
10	328	11.3	328	15	US-10-146-502-96
11	276	9.5	445	16	US-10-062-674-914
12	250	8.6	491	10	US-09-918-995-24717
13	216.8	7.4	302	12	US-10-152-319A-292
14	154.2	5.3	381	16	US-10-305-720-649

15	135.6	4.7	2972	13	US-10-342-887-1302	Sequence 1302, Ap
16	135.6	4.7	2972	13	US-10-260-708-24	Sequence 24, Ap1
17	135.6	4.7	2972	13	US-10-172-118-1302	Sequence 1302, Ap
18	135.6	4.7	2972	15	US-10-269-909-72	Sequence 72, Ap1
19	135.6	4.7	2972	16	US-10-159-563-102	Sequence 102, Ap1
20	134.4	4.6	2770	10	US-09-769-970-8	Sequence 8, Ap1
21	134.4	4.6	2783	15	US-10-024-238A-102	Sequence 102, Ap1
22	134.4	4.6	2783	15	US-10-042-211A-102	Sequence 102, Ap1
23	134.4	4.6	2783	17	US-10-617-217A-102	Sequence 102, Ap1
24	134.4	4.6	2783	16	US-10-264-049-949	Sequence 949, Ap1
25	134.4	4.6	3024	15	US-10-198-846-11031	Sequence 11031, A
26	133.4	4.6	2781	12	US-10-152-319A-1953	Sequence 1953, Ap
27	133.4	4.6	2781	16	US-10-191-803-222	Sequence 222, Ap1
28	131.4	4.5	2169	15	US-10-108-580-15	Sequence 15, Ap1
29	131.4	4.5	2169	15	US-10-204-041-15	Sequence 15, Ap1
30	131.4	4.5	2169	16	US-10-305-720-1147	Sequence 1147, Ap
31	131.4	4.5	2169	17	US-10-304-116-4	Sequence 4, Ap1
32	131.4	4.5	2169	17	US-10-620-052A-27	Sequence 27, Ap1
33	129.8	4.5	2378	17	US-10-755-889-3	Sequence 3, Ap1
34	128.2	4.4	2410	17	US-10-304-116-12	Sequence 12, Ap1
35	124	4.3	2278	15	US-10-059-585-13	Sequence 13, Ap1
36	122	4.2	1104	16	US-10-369-493-46146	Sequence 46146, A
37	121.8	4.2	2118	16	US-10-369-493-2551	Sequence 2551, A
38	118	4.1	2033	13	US-10-342-887-881	Sequence 881, Ap1
39	118	4.1	2033	13	US-10-172-118-881	Sequence 881, Ap1
40	115.2	4.0	1539	9	US-09-938-842A-162	Sequence 162, Ap1
41	115.2	4.0	1539	11	US-10-429-849-3	Sequence 3, Ap1
42	113.6	3.9	1815	13	US-10-429-849-3	Sequence 3, Ap1
43	113.2	3.9	2070	16	US-10-264-049-915	Sequence 915, Ap1
44	113.2	3.9	2146	14	US-10-071-766-83	Sequence 83, Ap1
45	113.2	3.9	2198	9	US-09-012-135A-2	Sequence 2, Ap1

## ALIGNMENTS

RESULT 1  
US-10-026-021-1  
; Sequence 1, Application US/10026021  
; Publication No. US20030027756A1  
; GENERAL INFORMATION:  
; APPLICANT: Hitoshi, Yasumichi  
; APPLICANT: Demo, Susan  
; APPLICANT: Jenkins, Yonchu  
; APPLICANT: Rigel Pharmaceuticals, Inc.  
; TITLE OF INVENTION: Treatment of Cellular Proliferation for  
; FILE REFERENCE: 021044-001210US  
; CURRENT APPLICATION NUMBER: US/10/026,021  
; CURRENT FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/309,632  
; PRIOR FILING DATE: 2001-08-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2913  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2913)  
; OTHER INFORMATION: human SAK serine/threonine kinase  
US-10-026-021-1

Query Match 100.0%; Score 2913; DB 15; Length 2913;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2913; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGGACCTGATCGGAGGAGATTCGAGTTTAAAGTTGGAATCTGCTGTAA 60  
DB 1 ATCGGACCTGATCGGAGGAGATTCGAGTTTAAAGTTGGAATCTGCTGTAA 60  
QY 61 GATCATTTGCTGTGTACAGAGCTGATTCACACTGTGGAAGTTGCATTC 120



```
Db      61  |||||
      61  GGAATCTTGGCTGGTGTCTACAGAGCTGATGCCATTCACACTGGTTTGGAGTTGCATC 120
Qy      121  AAAATGATGATTAAGAAAGCAATGACAAACGAGAAATGATACAGAGTCCAAATGAG 180
Db      121  AAAATGATGATTAAGAAAGCAATGACAAACGAGAAATGATACAGAGTCCAAATGAG 180
Qy      181  GTGAAATACATTCGCAATGAAACATCTCTATCTTGGAGCTTTTAACTATTTTGA 240
Db      181  GTGAAATACATTCGCAATGAAACATCTCTATCTTGGAGCTTTTAACTATTTTGA 240
Qy      241  GATGCAATTAATGATGATCTGATTAAGAAATGTCCTAATGAGAAATGAAACAGGTAT 300
Db      241  GATGCAATTAATGATGATCTGATTAAGAAATGTCCTAATGAGAAATGAAACAGGTAT 300
Qy      301  CTAAGAATAGAGTAAACCTTCTCAGAAAATGAGCTGACACTTCATGACCGAGATC 360
Db      301  CTAAGAATAGAGTAAACCTTCTCAGAAAATGAGCTGACACTTCATGACCGAGATC 360
Qy      361  ATCAAGAGGATGTTGATCTTCATTCATGATGATTAATACACCGGGAAGCTCAGACTTCT 420
Db      361  ATCAAGAGGATGTTGATCTTCATTCATGATGATTAATACACCGGGAAGCTCAGACTTCT 420
Qy      421  AACCTCTACTGACTCGTAATATGAAACATGAAATGCTGATTTTGGGCTGGCACTCA 480
Db      421  AACCTCTACTGACTCGTAATATGAAACATGAAATGCTGATTTTGGGCTGGCACTCA 480
Qy      481  CTGAAAATGCGACATGAAAAGCACTATACATTAATGAGAACTCTAATCAATTTCCACA 540
Db      481  CTGAAAATGCGACATGAAAAGCACTATACATTAATGAGAACTCTAATCAATTTCCACA 540
Qy      541  GAAATGGCGACTGGAAGTGCACATGGCTTGAATCTGATGTTTGGTCCCTGGGCTGTATG 600
Db      541  GAAATGGCGACTGGAAGTGCACATGGCTTGAATCTGATGTTTGGTCCCTGGGCTGTATG 600
Qy      601  TTTTATACATTAATTAATGCGAGAACCACTTCGACACTGACACAGTCAAGAACATTA 660
Db      601  TTTTATACATTAATTAATGCGAGAACCACTTCGACACTGACACAGTCAAGAACATTA 660
Qy      661  AATTAAGTATGATGCGAATTAATGAAATGCCATCTTTTGTCAATAGAGGCCAAGAC 720
Db      661  AATTAAGTATGATGCGAATTAATGAAATGCCATCTTTTGTCAATAGAGGCCAAGAC 720
Qy      721  CTATATCACCAAGTACTTGTGTAAGAAATCGACAGATCGTTAAGTCTGTCAGATTTG 780
Db      721  CTATATCACCAAGTACTTGTGTAAGAAATCGACAGATCGTTAAGTCTGTCAGATTTG 780
Qy      781  GACCATCTTTTATGTCCTCGAAATCTTCAACAAAAGTAAAGATTAGAACTGTGAA 840
Db      781  GACCATCTTTTATGTCCTCGAAATCTTCAACAAAAGTAAAGATTAGAACTGTGAA 840
Qy      841  GACTCAATGATAGTGGGATGCGCAATTTCTACTGCAATTAAGCTTTCTTCCAGTACC 900
Db      841  GACTCAATGATAGTGGGATGCGCAATTTCTACTGCAATTAAGCTTTCTTCCAGTACC 900
Qy      901  AGTATAGTGTAGTATTATTTGACAAAAGAGACTTTTGATTTGTCAGCCATCCCAAT 960
Db      901  AGTATAGTGTAGTATTATTTGACAAAAGAGACTTTTGATTTGTCAGCCATCCCAAT 960
Qy      961  AAAATGACCTGATTTTCCAAAAGAAATTAAGTTCACTGATTTTCTTTCAGAGATGGA 1020
Db      961  AAAATGACCTGATTTTCCAAAAGAAATTAAGTTCACTGATTTTCTTTCAGAGATGGA 1020
Qy      1021  AACAGTTTATTAATCACTGATGGGAAATCAAGAAACAGTAAATAGTGAAGGGAAGGTA 1080
Db      1021  AACAGTTTATTAATCACTGATGGGAAATCAAGAAACAGTAAATAGTGAAGGGAAGGTA 1080
Qy      1081  AATCAAGATGCAAGAAAGGCCACATTCGATACCTTCGTAGAGCTTATTCCTCTGAT 1140
Db      1081  AATCAAGATGCAAGAAAGGCCACATTCGATACCTTCGTAGAGCTTATTCCTCTGAT 1140
Qy      1141  AGATCGGCACTTCTAATAGTCAGTTCACGAAACATTAATAGGAACGATGTCAC 1200
Db      1141  AGATCTGGCACTTCTAATAGTCAGTTCACGAAACATTAATAGGAACGATGTCAC 1200

Db      1141  AGATCTGGCACTTCTAATAGTCAGTTCACGAAACATTAATAGGAACGATGTCAC 1200
Qy      1201  TCAGCAAAATGCTTTCACTGTCCAAAAGATCAGAGGAGGTGAAAATGAAAGAGGTAC 1260
Db      1201  TCAGCAAAATGCTTTCACTGTCCAAAAGATCAGAGGAGGTGAAAATGAAAGAGGTAC 1260
Qy      1261  TCACCCACAGACACATGTCGCAACTTTTAACTTTTAAAGAAAAGACATCCAGTAGT 1320
Db      1261  TCACCCACAGACACATGTCGCAACTTTTAACTTTTAAAGAAAAGACATCCAGTAGT 1320
Qy      1321  TCTGATCTTTTGAAGAAGCTGATTAACATCAAGACCTCCAAATCACTTTGTCAGAGA 1380
Db      1321  TCTGATCTTTTGAAGAAGCTGATTAACATCAAGACCTCCAAATCACTTTGTCAGAGA 1380
Qy      1381  AAAATCTCTTTTCCATTTTGACAGACCCGACACTCGACTGAACCCGTAAACAGTGTAT 1440
Db      1381  AAAATCTCTTTTCCATTTTGACAGACCCGACACTCGACTGAACCCGTAAACAGTGTAT 1440
Qy      1441  GGGAAATGCAAAATTAAGTCACTTTAAGAAAATCTAATTAATGACAGATCAAGCCA 1500
Db      1441  GGGAAATGCAAAATTAAGTCACTTTAAGAAAATCTAATTAATGACAGATCAAGCCA 1500
Qy      1501  AACCGGGAATTCAGAGGCAATCCAGATTTGCAAGAGACATCAAAAAAATGCTGAGACT 1560
Db      1501  AACCGGGAATTCAGAGGCAATCCAGATTTGCAAGAGACATCAAAAAAATGCTGAGACT 1560
Qy      1561  GATTCAAAATGCAAAAAAGAACTCTGATGCTTGTATATGCAATTCGTAAAAACAGCAA 1620
Db      1561  GATTCAAAATGCAAAAAAGAACTCTGATGCTTGTATATGCAATTCGTAAAAACAGCAA 1620
Qy      1621  AATACATGAATATATAGCTGCACTTCAAGTAACCTGAGATTAATCCAAAGAAATGT 1680
Db      1621  AATACATGAATATATAGCTGCACTTCAAGTAACCTGAGATTAATCCAAAGAAATGT 1680
Qy      1681  GTTTTGGCTGATTCCTCTTTCTGAAACAGAGCAAGCTAGGGGTATGAGCCACATGG 1740
Db      1681  GTTTTGGCTGATTCCTCTTTCTGAAACAGAGCAAGCTAGGGGTATGAGCCACATGG 1740
Qy      1741  GGTATACGAATGATACATTAAGAAAGATCAATCCCGTTGTGTGTCTCAAGTTAAAA 1800
Db      1741  GGTATACGAATGATACATTAAGAAAGATCAATCCCGTTGTGTGTCTCAAGTTAAAA 1800
Qy      1801  CCAATCAAGACAAAACCAAAAAAGCTGTGAGATCACTTGATTCAGAGAGGTGTG 1860
Db      1801  CCAATCAAGACAAAACCAAAAAAGCTGTGAGATCACTTGATTCAGAGAGGTGTG 1860
Qy      1861  GTGAGCTGTGTAAGAGATGATCTCAAGAAATATGTGAAGAAAGTTCTTCAGATATCT 1920
Db      1861  GTGAGCTGTGTAAGAGATGATCTCAAGAAATATGTGAAGAAAGTTCTTCAGATATCT 1920
Qy      1921  AGTATGAAATATGATCACTATTTATTAATCCAAATGATGTGATGAGTTCCTCTTGCT 1980
Db      1921  AGTATGAAATATGATCACTATTTATTAATCCAAATGATGTGATGAGTTCCTCTTGCT 1980
Qy      1981  GATGACCAACCTTCACCTACTGACAAATCAATCAAGTATGAGTTCGATTAACAGAA 2040
Db      1981  GATGACCAACCTTCACCTACTGACAAATCAATCAAGTATGAGTTCGATTAACAGAA 2040
Qy      2041  AAATACGTGGGAAAATATCAATATGCTTCAGGTTTGTACAGTTCGATTAATATCT 2100
Db      2041  AAATACGTGGGAAAATATCAATATGCTTCAGGTTTGTACAGTTCGATTAATATCT 2100
Qy      2101  CCCAAATCACTTATTTTCAAGATATGCTTAATGATGATTTGATGAGAGATTCCTGCT 2160
Db      2101  CCCAAATCACTTATTTTCAAGATATGCTTAATGATGATTTGATGAGAGATTCCTGCT 2160
Qy      2161  GCTGATTTGAGTGTGTTTATGATGAGGTTTAAATACAAAACAGAAATTCATTT 2220
Db      2161  GCTGATTTGAGTGTGTTTATGATGAGGTTTAAATACAAAACAGAAATTCATTT 2220
Qy      2221  CAGGTATTTGAAAAGACAGGGAAGTCTTAACATTTAAAGGAAAGGATTAATGAC 2280
Db      2221  CAGGTATTTGAAAAGACAGGGAAGTCTTAACATTTAAAGGAAAGGATTAATGAC 2280
```

```
QY 2281 TTGAAGAGGAGATATAAATGTATATGACCATGCTAATAGAGGTGATGCTATTTGTTTA 2340
Db 2281 TTGAAGAGGAGATATAAATGTATATGACCATGCTAATAGAGGTGATGCTATTTGTTTA 2340
QY 2341 GCACGTGAATCCATAATTTTCAAGAGAGAAAGAACTAGAGTGTCTCCCTTTTCCCA 2400
Db 2341 GCACGTGAATCCATAATTTTCAAGAGAGAAAGAACTAGAGTGTCTCCCTTTTCCCA 2400
QY 2401 ATATATATAGAAAGAAACCTGTAGTACTAGTTCACTTAAGCCCTTATCACTCTCTCT 2460
Db 2401 ATATATATAGAAAGAAACCTGTAGTACTAGTTCACTTAAGCCCTTATCACTCTCTCT 2460
QY 2461 TCTGTGATTTCAATTTACCAAGAGAGATAGAGCATCTTTCAACGAATGTCATGCAAT 2520
Db 2461 TCTGTGATTTCAATTTACCAAGAGAGATAGAGCATCTTTCAACGAATGTCATGCAAT 2520
QY 2521 AGTGTCTTCTTCCACACAGGACCAATCTTAATCCCTTAATGTTTCAAAATGAAGA 2580
Db 2521 AGTGTCTTCTTCCACACAGGACCAATCTTAATCCCTTAATGTTTCAAAATGAAGA 2580
QY 2581 CTGGTCTTCAACATCAAGCTTTGGAACAGACATCTCTTAATAGTCTAAAGATTGT 2640
Db 2581 CTGGTCTTCAACATCAAGCTTTGGAACAGACATCTCTTAATAGTCTAAAGATTGT 2640
QY 2641 CTTCCTAATTCAGCAACCTTTGAAATCTGTTTGTGAAAAATGTTGGTGGCTACA 2700
Db 2641 CTTCCTAATTCAGCAACCTTTGAAATCTGTTTGTGAAAAATGTTGGTGGCTACA 2700
QY 2701 CAGTTAAGTGTGAGCTGTGTGGTTCAGTTTAATGATGGTCCAGTTGTTGTCAG 2760
Db 2701 CAGTTAAGTGTGAGCTGTGTGGTTCAGTTTAATGATGGTCCAGTTGTTGTCAG 2760
QY 2761 GCAGAGGTCTTCTTATCAGTTATACCTCAACAAATGTCAAACATAGATGAGAA 2820
Db 2761 GCAGAGGTCTTCTTATCAGTTATACCTCAACAAATGTCAAACATAGATGAGAA 2820
QY 2821 AATGAAATTTACCAACATCAATCAACAAAGAAATTCAGTGTCTTCCATCCCTTTG 2880
Db 2821 AATGAAATTTACCAACATCAATCAACAAAGAAATTCAGTGTCTTCCATCCCTTTG 2880
QY 2881 ATGTTTCTAATCCGACTCTTAATTTTCAATGA 2913
Db 2881 ATGTTTCTAATCCGACTCTTAATTTTCAATGA 2913

RESULT 2
US-10-037-270-280
Sequence 280, Application US/10037270
Publication No. US20030104529A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Aundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aifeng J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhimei
APPLICANT: Tiliinghaer, John
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/10/037,270
CURRENT FILING DATE: 2002-01-04
```

```

; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_genes Version 1.0
; SEQ ID NO 280
; LENGTH: 3937
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (334) ..(3246)
US-10-037-270-280

Query Match 99.9%; Score 2911.4; DB 15; Length 3937;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCTGCATCGGGAGAGATCGAGATTTTAAAGTTGAAATCGCTGTGATAA 60
Db 334 ATGGCGACCTGCATCGGGAGAGATCGAGATTTTAAAGTTGAAATCGCTGTGATAA 393
QY 61 GGATCATTTGCTGTGTCTACAGAGCTGATTCATTCACTGTGTTGAAAGTTGCAATC 120
Db 394 GGATCATTTGCTGTGTCTACAGAGCTGATTCATTCACTGTGTTGAAAGTTGCAATC 453
QY 121 AAATATATGATTAAGAAACCATGTTCAAAGCAGAAATGCTACAGAGTCCAAATGAG 180
Db 454 AAATATATGATTAAGAAACCATGTTCAAAGCAGAAATGCTACAGAGTCCAAATGAG 513
QY 181 GTGAAATATCATTTGCCAATTTGAAACATCTCTCATCTTGAGAGTTTAACTATTGAA 240
Db 514 GTGAAATATCATTTGCCAATTTGAAACATCTCTCATCTTGAGAGTTTAACTATTGAA 573
QY 241 GATAGCAATTAATGTATCTGTATTTAGAAATGTGCAATATGAGAAATGAAACAGTAT 300
Db 574 GATAGCAATTAATGTATCTGTATTTAGAAATGTGCAATATGAGAAATGAAACAGTAT 633
QY 301 CTAAAGATAGAGTGAACCCCTTCTCAAAATGAAAGCTGACATTCATGACCGAGATC 360
Db 634 CTAAAGATAGAGTGAACCCCTTCTCAAAATGAAAGCTGACATTCATGACCGAGATC 693
QY 361 ATCACAGGATGTGTATCTTCAATCTCATGTATCTACACCGGAGCTCACACTTCT 420
Db 694 ATCACAGGATGTGTATCTTCAATCTCATGTATCTACACCGGAGCTCACACTTCT 753
QY 421 AACCTCTACTGACTGTATATGAACATCAAGATTGCTGATTTGGCTGCACTCAA 480
Db 754 AACCTCTACTGACTGTATATGAACATCAAGATTGCTGATTTGGCTGCACTCAA 813
QY 481 CTGAAATATCCCATGAAAGCACTATACATTAATGTGAACTCTTCACTCAATTTGACA 540
Db 814 CTGAAATATCCCATGAAAGCACTATACATTAATGTGAACTCTTCACTCAATTTGACA 873
QY 541 GAAATGCCACTGCAAGTGCACATGGCTTGAATCGATGTTTGGCTGGCTGTATG 600
Db 874 GAAATGCCACTGCAAGTGCACATGGCTTGAATCGATGTTTGGCTGGCTGTATG 933
QY 601 TTTTATACATTAATCTTATCGGAGACACACCTTGCACATGACAGTCAAGAACATTTA 660
Db 934 TTTTATACATTAATCTTATCGGAGACACACCTTGCACATGACAGTCAAGAACATTTA 993
QY 661 AATAAAGTATATGGCAGATTAATGAATGCCATCTTTTGTCAATAGAGCCAAAGAC 720
Db 994 AATAAAGTATATGGCAGATTAATGAATGCCAACTTTTGTCAATAGAGCCAAAGAC 1053
QY 721 CTATATCACAGTACTTGTGTAAGATCGAGAGATCGTTAAGTGTCTTCAAGATG 780
Db 1054 CTATATCACAGTACTTGTGTAAGATCGAGAGATCGTTAAGTGTCTTCAAGATG 1113
QY 781 GACCATCTTTTATGTCCGAAATCTTTCACAAAGTAAAGATTAGAACTGTGAA 840
```

Db 1114 GACCACTCTTTATGATGCCGAATCTTCAACAAAAGTAAGATTAGGAAGCTGTGAA 1173  
 Qy 841 GACTCAATTGATAGGGCAATGCCAATTCTTCTGCAATACAGCTTCTTCAAGTACC 900  
 Db 1174 GACTCAATTGATAGGGCAATGCCAATTCTTCTGCAATACAGCTTCTTCAAGTACC 1233  
 Qy 901 AGATAAGTGTAGTATTATTTGACAAAAGAACCTTTGATGGTCAAGCACTCCCAAT 960  
 Db 1234 AGATAAGTGTAGTATTATTTGACAAAAGAACCTTTGATGGTCAAGCACTCCCAAT 1293  
 Qy 961 AAAATGCTGATTTCCAAAAGATTAAGTTCACTGATTTTCTTCTTCAAGATGGA 1020  
 Db 1294 AAAATGCTGATTTCCAAAAGATTAAGTTCACTGATTTTCTTCTTCAAGATGGA 1353  
 Qy 1021 AACGTTTTTATCTAGTGGGGAATCAAGAAACGATAATGTGGAGGGGAAGATGA 1080  
 Db 1354 AACGTTTTTATCTAGTGGGGAATCAAGAAACGATAATGTGGAGGGGAAGATGA 1413  
 Qy 1081 ATTCAAGATCAGAGAAAGGCCACATTCGTGATACCTTGTGAGGCTTAATTCCTGTAT 1140  
 Db 1414 ATTCAAGATCAGAGAAAGGCCACATTCGTGATACCTTGTGAGGCTTAATTCCTGTAT 1473  
 Qy 1141 AGATCGGCACTCTATATAGCTCTCAAGCAAAAACATATCATATGAGACATGTCAAC 1200  
 Db 1474 AGATCGGCACTCTATATAGCTCTCAAGCAAAAACATATCATATGAGACATGTCAAC 1533  
 Qy 1201 TCAGCAAAATGCTTTCAGTGTCCAAAAGATCAGAGAGGTGAATGAGAGAGTAC 1260  
 Db 1534 TCAGCAAAATGCTTTCAGTGTCCAAAAGATCAGAGAGGTGAATGAGAGAGTAC 1593  
 Qy 1261 TCACCCACAGACAAACATGCGCAATTTTAACTTTTAAAGAAAAGACATCCAGTAGT 1320  
 Db 1594 TCACCCACAGACAAACATGCGCAATTTTAACTTTTAAAGAAAAGACATCCAGTAGT 1653  
 Qy 1321 TCTGATCTTTTGAAGAAGCTGATTAACAATCAAGCACTCTCAATATCTTTGTCAGAGA 1380  
 Db 1654 TCTGATCTTTTGAAGAAGCTGATTAACAATCAAGCACTCTCAATATCTTTGTCAGAGA 1713  
 Qy 1381 AAAATCTCTTTTCATTTGCAAGCCGACACCTCAGACTGAAAACCGTACACAGTGGTTT 1440  
 Db 1714 AAAATCTCTTTTCATTTGCAAGCCGACACCTCAGACTGAAAACCGTACAGTGGTTT 1773  
 Qy 1441 GGGAACTGTCAAAATTAATGCTCATTTTAAAGAAAACATCTGAATATGACAGCATCAGCCCA 1500  
 Db 1774 GGGAACTGTCAAAATTAATGCTCATTTTAAAGAAAACATCTGAATATGACAGCATCAGCCCA 1833  
 Qy 1501 AACCGGACTTCAGGGCCATCAGATTTGCAAGAGGACACATCAAAAATATCCCTGAGACT 1560  
 Db 1834 AACCGGACTTCAGGGCCATCAGATTTGCAAGAGGACACATCAAAAATATCCCTGAGACT 1893  
 Qy 1561 GATACAAAAGTCAAAAAGAACTCTGATGCTTGTATATGCAACTTCTGTAAACAGCA 1620  
 Db 1894 GATACAAAAGTCAAAAAGAACTCTGATGCTTGTATATGCAACTTCTGTAAACAGCA 1953  
 Qy 1621 AATACCATGAATATATATGCTGCACTTCAAGTAAACCTGAGATTAATCCAAAGAAATGT 1680  
 Db 1954 AATACCATGAATATATATGCTGCACTTCAAGTAAACCTGAGATTAATCCAAAGAAATGT 2013  
 Qy 1681 GTTTTGGCTCAGATCTCTTCTTCTGAAACAGACAGACCTAGGGGTATGGAAGCACAAGG 1740  
 Db 2014 GTTTTGGCTCAGATCTCTTCTTCTGAAACAGACAGACCTAGGGGTATGGAAGCACAAGG 2073  
 Qy 1741 GGTATACGAATCGATACATTAAGAGCATATCTCCGTTGGTGTCTCAACAGTTAAAA 1800  
 Db 2074 GGTATACGAATCGATACATTAAGAGCATATCTCCGTTGGTGTCTCAACAGTTAAAA 2133  
 Qy 1801 CCAATCAGACAGAAAACCAAAAAGGCTGTGTGAGCATATCTGATTCAAGAGAGGTGTGT 1860  
 Db 2134 CCAATCAGACAGAAAACCAAAAAGGCTGTGTGAGCATATCTGATTCAAGAGAGGTGTGT 2193  
 Qy 1861 GTGAGCTTTGTAAAGAGATGATCATCTCAAGAAATATGTGAAGAAAGTTCTTCAAGATATCT 1920  
 Db 2194 GTGAGCTTTGTAAAGAGATGATCATCTCAAGAAATATGTGAAGAAAGTTCTTCAAGATATCT 2253

Qy 1921 AGTAGAGAAATACATACATATTTATATCCAAAATGCTGTAGAGTTTCTCTGTGCT 1980  
 Db 2254 AGTAGAGAAATACATACATATTTATATCCAAAATGCTGTAGAGTTTCTCTGTGCT 2313  
 Qy 1981 GATPAGCCAGCCCTCACCTATGACAAACATCACTAGTATGACAGTTTGAACATTTACAGAA 2040  
 Db 2314 GATPAGCCAGCCCTCACCTATGACAAACATCACTAGTATGACAGTTTGAACATTTACAGAA 2373  
 Qy 2041 AAATACCTGCGAAAATATCAATATGCTTCCAGTTTGTATACAGCTTGAATCTAAATCT 2100  
 Db 2374 AAATACCTGCGAAAATATCAATATGCTTCCAGTTTGTATACAGCTTGAATCTAAATCT 2433  
 Qy 2101 CCCAAAATCATCTATTTTACAAAGATATGCTAAATGCAATTTTGTAGAGAAATTCCTGCT 2160  
 Db 2434 CCCAAAATCATCTATTTTACAAAGATATGCTAAATGCAATTTTGTAGAGAAATTCCTGCT 2493  
 Qy 2161 GCTGATTTTGAAGTTTGGTTTTATGATGAGGGTAAATATACAAAGAGAAAGTTTCATT 2220  
 Db 2494 GCTGATTTTGAAGTTTGGTTTTATGATGAGGGTAAATATACAAAGAGAAAGTTTCATT 2553  
 Qy 2221 CAGGTGATTTGAAAAGACAGGGAAGTCTTACACTTTAAAGTGAAGTGAATTAATAGC 2280  
 Db 2554 CAGGTGATTTGAAAAGACAGGGAAGTCTTACACTTTAAAGTGAAGTGAATTAATAGC 2613  
 Qy 2281 TTGAAGAGGAGATPAAAATGATATATGGAACAATGCTATGAGGGTCACTGATTTGTTTA 2340  
 Db 2614 TTGAAGAGGAGATPAAAATGATATATGGAACAATGCTATGAGGGTCACTGATTTGTTTA 2673  
 Qy 2341 GCACTGGAATCATATATTTCAAGAGAGAAAGAAAACCTAGAGTGTCTCCCTTTTCCCA 2400  
 Db 2674 GCACTGGAATCATATATTTCAAGAGAGAAAGAAAACCTAGAGTGTCTCCCTTTTCCCA 2733  
 Qy 2401 AATAATCATAAGAAAGAAAACCTGTGATATCTTCACTTAAAGCTTATCACTCTCTCT 2460  
 Db 2734 AATAATCATAAGAAAGAAAACCTGTGATATCTTCACTTAAAGCTTATCACTCTCTCT 2793  
 Qy 2461 TCTGTGATTCAAATTTCCCAAGAGAGATGAGCATCTTTCAACAGAAATGCTATGAT 2520  
 Db 2794 TCTGTGATTCAAATTTCCCAAGAGAGATGAGCATCTTTCAACAGAAATGCTATGAT 2853  
 Qy 2521 AGTGTGCTTCTCCAAACAGGCAACCAATCCTTATATCTTATGTTTAAACATGAAGGA 2580  
 Db 2854 AGTGTGCTTCTCCAAACAGGCAACCAATCCTTATATCTTATGTTTAAACATGAAGGA 2913  
 Qy 2581 CTTGGTCTTAACTAATCAGCTTGTGAAACAGACATCTCTTAAATATCTTAAATGAATGT 2640  
 Db 2914 CTTGGTCTTAACTAATCAGCTTGTGAAACAGACATCTCTTAAATATCTTAAATGAATGT 2973  
 Qy 2641 CTTCTAATCAGACAACTTTTGAAGATCTGTTTTGTGAAGAAATGTTGGTGGCTACA 2700  
 Db 2974 CTTCTAATCAGACAACTTTTGAAGATCTGTTTTGTGAAGAAATGTTGGTGGCTACA 3033  
 Qy 2701 CAGTTTAACTATGAGAGCTGTGTGGGTTCAAGTTTATGATGAGTCCAGTTGGTTGACAG 2760  
 Db 3034 CAGTTTAACTATGAGAGCTGTGTGGGTTCAAGTTTATGATGAGTCCAGTTGGTTGACAG 3093  
 Qy 2761 GCAGAGTGTCTTATCAGTTTAACTCTCAACAAATGCTCAAAACACTAGATATGAGAA 2820  
 Db 3094 GCAGAGTGTCTTATCAGTTTAACTCTCAACAAATGCTCAAAACACTAGATATGAGAA 3153  
 Qy 2821 AATGAAAATTTACAGACTACATCAACAGAAATTAACAGTGTGTCTTCACTCTTTTG 2880  
 Db 3154 AATGAAAATTTACAGACTACATCAACAGAAATTAACAGTGTGTCTTCACTCTTTTG 3213  
 Qy 2881 ATGTTTTCTAATCGAATCTCTAATTTTCAATGA 2913  
 Db 3214 ATGTTTTCTAATCGAATCTCTAATTTTCAATGA 3246

RESULT 3  
 US-10-117-722-280  
 ; Sequence 280, Application US/10117722

Publication No. US20030219744A1  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Abundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and  
FILE REFERENCE: 784CIP2BCTP  
CURRENT APPLICATION NUMBER: US/10/117,722  
CURRENT FILING DATE: 2002-04-04  
PRIOR APPLICATION NUMBER: 09/520,312  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1104  
SOFTWARE: pl\_Fl\_genes Version 1.0  
SEQ ID NO 280  
LENGTH: 3937  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (334)..(3246)  
US-10-117-722-280

Query Match 99.9%; Score 2911.4; DB 16; Length 3937;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCACTGTCATCGGGGAGAAATCGAGATTTTAAAGTTGGAATCTGCTGTAA 60  
DB 334 ATGGGCACTGTCATCGGGGAGAAATCGAGATTTTAAAGTTGGAATCTGCTGTAA 393  
QY 61 GGATCATTTTCTGCTGCTCAAGAGCTGATCCATTCACACGTGGTTGGAAGTTGCAATC 120  
DB 394 GGATCATTTTCTGCTGCTCAAGAGCTGATCCATTCACACGTGGTTGGAAGTTGCAATC 453  
QY 121 AAAATGATGATTAAGAAAGCAATGTACAAAGCAGAAATGTAACAGAGTCCAAATGAG 180  
DB 454 AAAATGATGATTAAGAAAGCAATGTACAAAGCAGAAATGTAACAGAGTCCAAATGAG 513  
QY 181 GTGAAATATCATTTGCCAATTGAAATCTCTTCTATCTTGGAGCTTATACTATTTTGA 240  
DB 514 GTGAAATATCATTTGCCAATTGAAATCTCTTCTATCTTGGAGCTTATACTATTTTGA 573  
QY 241 GATTAAGATTTATGTGTATCTGTATTAAGAAATGTCATTAAGAAATGAAACAGTAT 300  
DB 574 GATTAAGATTTATGTGTATCTGTATTAAGAAATGTCATTAAGAAATGAAACAGTAT 633  
QY 301 CTAAAGATAGAGTAAACCTTCTCAGAAATGAAGTGCACACTTCATGACAGATC 360  
DB 634 CTAAAGATAGAGTAAACCTTCTCAGAAATGAAGTGCACACTTCATGACAGATC 693  
QY 361 ATCAAGGAGTGTGTATCTTCAATTCATGTATACACCGGAGCTCACAATTTCT 420  
DB 694 ATCAAGGAGTGTGTATCTTCAATTCATGTATACACCGGAGCTCACAATTTCT 753  
QY 421 AACCTCTCTCTGATCTGTATTAAGAAATCAAGATTTGTTGGCTGGGAACTCAA 480  
DB 754 AACCTCTCTCTGATCTGTATTAAGAAATCAAGATTTGTTGGCTGGGAACTCAA 813  
QY 481 CTGAAATGCAATGAAAGCACTATTAATTAATGAACTCTTAATCAATTTGACCA 540  
DB 814 CTGAAATGCAATGAAAGCACTATTAATTAATGAACTCTTAATCAATTTGACCA 873  
QY 541 GAAATGCACTGCAAGTGCACATGACCTTGAATCTGATGTTTGGTCCCTGGCTGTATG 600  
DB 874 GAAATGCACTGCAAGTGCACATGACCTTGAATCTGATGTTTGGTCCCTGGCTGTATG 933

QY 601 TTTTATACATTAATCTATCGGAGACCACTTTCGACACTGACAGCTCAAGACATTA 660  
DB 934 TTTTATACATTAATCTATCGGAGACCACTTTCGACACTGACAGCTCAAGACATTA 993  
QY 661 AATTAAGTATGTTGGCAGATTTATGAATGCCATCTTTTGTCAATTAAGCCAAAGAC 720  
DB 994 AATTAAGTATGTTGGCAGATTTATGAATGCCAATCTTTTGTCAATTAAGCCAAAGAC 1053  
QY 721 CTTTATCACCAATTCCTGTGAAGAACGAGAGATCGTTTAAGTGTCTTCAGTATG 780  
DB 1054 CTTTATCACCAATTCCTGTGAAGAACGAGAGATCGTTTAAGTGTCTTCAGTATG 1113  
QY 781 GACATCTCTTATATGTCGGAATCTTCAACAAAGTAAAGATTAGAACTGTGAA 840  
DB 1114 GACATCTCTTATATGTCGGAATCTTCAACAAAGTAAAGATTAGAACTGTGAA 1173  
QY 841 GACTCAATTAAGTGGGATGCAATTTCTACTGCAATTAACGCTTCTTCAGTATC 900  
DB 1174 GACTCAATTAAGTGGGATGCAATTTCTACTGCAATTAACGCTTCTTCAGTATC 1233  
QY 901 AGTATTAAGTATGTTTATTAAGCAAAAGAACTTTGATTTGTCAGGCACTCCCAAT 960  
DB 1234 AGTATTAAGTATGTTTATTAAGCAAAAGAACTTTGATTTGTCAGGCACTCCCAAT 1293  
QY 961 AAAATGATGATTTTCCAAAGATTAAGTCACTGATTTTCTTCTCAGAGATGGA 1020  
DB 1294 AAAATGATGATTTTCCAAAGATTAAGTCACTGATTTTCTTCTCAGAGATGGA 1353  
QY 1021 AACAGTTTATTAATCAATGAGGAAATCAAGAAACCAATTAATGAGGAGGAAAGTA 1080  
DB 1354 AACAGTTTATTAATCAATGAGGAAATCAAGAAACCAATTAATGAGGAGGAAAGTA 1413  
QY 1081 ATTCAAGATGCAAGAAAGGCAATCTGATCTCTGATCTCTGATCTCTGAT 1140  
DB 1414 ATTCAAGATGCAAGAAAGGCAATCTGATCTCTGATCTCTGATCTCTGAT 1473  
QY 1141 AGATCTGCACTTCTTAATGATGATCTCAAGCAAAACATATATGAAAGATGCA 1200  
DB 1474 AGATCTGCACTTCTTAATGATGATCTCAAGCAAAACATATATGAAAGATGCA 1533  
QY 1201 TCAGCAAAATGCTTCAATGATCTCAAGCAAAACATATATGAAAGATGCA 1260  
DB 1534 TCAGCAAAATGCTTCAATGATCTCAAGCAAAACATATATGAAAGATGCA 1593  
QY 1261 TCAGCAAAATGCTTCAATGATCTCAAGCAAAACATATATGAAAGATGCA 1320  
DB 1594 TCAGCAAAATGCTTCAATGATCTCAAGCAAAACATATATGAAAGATGCA 1653  
QY 1321 TCTGATCTTTTGAAGAACTGATTAACATCAAGCACTCTCAATCATCTTTGCCAGA 1380  
DB 1654 TCTGATCTTTTGAAGAACTGATTAACATCAAGCACTCTCAATCATCTTTGCCAGA 1713  
QY 1381 AAAATCTCTTTCATTTGCAATGCAAGCCGACCTCAGACTGAAACCGTCAACAGTGT 1440  
DB 1714 AAAATCTCTTTCATTTGCAATGCAAGCCGACCTCAGACTGAAACCGTCAACAGTGT 1773  
QY 1441 GGAATCTGCAATTAATGCTCAATTAAGAAACATCACTGATTAAGCAAGCATGCGCCA 1500  
DB 1774 GGAATCTGCAATTAATGCTCAATTAAGAAACATCACTGATTAAGCAAGCATGCGCCA 1833  
QY 1501 AACCGGACCTTCAAGGACCTCAGATTTGCAAGAAAGCACTCAAAATAATGCTGGAAT 1560  
DB 1834 AACCGGACCTTCAAGGACCTCAGATTTGCAAGAAAGCACTCAAAATAATGCTGGAAT 1893  
QY 1561 GATTAAGATGCAAAAGAACTCTGATGTTCTGATTAATGCAATTTCTGTAACAGCAA 1620  
DB 1894 GATTAAGATGCAAAAGAACTCTGATGTTCTGATTAATGCAATTTCTGTAACAGCAA 1953  
QY 1621 AATTAAGATTAATTAATGCACTTCACTTAACCTGAAATTAATTAATTAATTAAT 1680  
DB 1954 AATTAAGATTAATTAATGCACTTCACTTAACCTGAAATTAATTAATTAATTAAT 2013  
QY 1681 GTTTTGGCTCAGATCTCTTCTGAAACAGAACTAAGGGGTATGAGGCCACCATGG 1740

Db 2014 GTTTTGGCTGAGTCTCTTTCTGAACAGCAAGACTGAGGGATATGAGCACCACATGG 2073  
 Qy 1741 GGTTCAGAAATCCGATCACTTAAGAAGCACTTACATCTCCGTTGGTCTCACAGGTTAAA 1800  
 Db 2074 GGTTCAGAAATCCGATCACTTAAGAAGCACTTACATCTCCGTTGGTCTCACAGGTTAAA 2133  
 Qy 1801 CCAATCAGACAGAAAACCAAAAAGGCTGTGGTGAACATCTGATTCAAGAGAGTGTGT 1860  
 Db 2134 CCAATCAGACAGAAAACCAAAAAGGCTGTGGTGAACATCTGATTCAAGAGAGTGTGT 2193  
 Qy 1861 GTGGAGCTTTGAAGAGATATGCACTTCAGAAATATGTGAAGAAGTCTTCAGATATCT 1920  
 Db 2194 GTGGAGCTTTGAAGAGATATGCACTTCAGAAATATGTGAAGAAGTCTTCAGATATCT 2253  
 Qy 1921 AGTATGGAATATGATCACTATTATTAATCAAAATGTGTAGAGGTTTCCCTTGTCT 1980  
 Db 2254 AGTATGGAATATGATCACTATTATTAATCAAAATGTGTAGAGGTTTCCCTTGTCT 2313  
 Qy 1981 GATGACCACTTCACCTACTGACACATCATGATGATGATGATGATGATGATGATGATGAT 2040  
 Db 2314 GATGACCACTTCACCTACTGACACATCATGATGATGATGATGATGATGATGATGATGAT 2373  
 Qy 2041 AAATACCTGGGGAATATCAATATGCTTCCAGGTTTGTACAGCTTGTATGATGATGATGAT 2100  
 Db 2374 AAATACCTGGGGAATATCAATATGCTTCCAGGTTTGTACAGCTTGTATGATGATGATGATGAT 2433  
 Qy 2101 CCCAAATCACTTAATTTTACAGATATGTAATGCACTTTGATGAGAAATCTCCGTGGT 2160  
 Db 2434 CCCAAATCACTTAATTTTACAGATATGTAATGCACTTTGATGAGAAATCTCCGTGGT 2493  
 Qy 2161 GCTGATTTTGAAGGTTGGTTTATGATGGGTTAAAATACAGAAAACAGAGATTTTCAT 2220  
 Db 2494 GCTGATTTTGAAGGTTGGTTTATGATGGGTTAAAATACAGAAAACAGAGATTTTCAT 2553  
 Qy 2221 CAGGTGATGAAAGAAGAGAGAGTCTTACACTTTTAAAAGGTAAGTAAATATAC 2280  
 Db 2554 CAGGTGATGAAAGAAGAGAGAGTCTTACACTTTTAAAAGGTAAGTAAATATAC 2613  
 Qy 2281 TTGAAAAGAGAGATTAATAATGATATGATGATGATGATGATGATGATGATGATGATGAT 2340  
 Db 2614 TTGAAAAGAGAGATTAATAATGATATGATGATGATGATGATGATGATGATGATGATGAT 2673  
 Qy 2341 GCACGTGAATCCATATTTCAAG 2400  
 Db 2674 GCACGTGAATCCATATTTCAAG 2733  
 Qy 2401 ATATATATGAG 2460  
 Db 2734 ATATATATGAG 2793  
 Qy 2461 TCTGTGAGATTCAAATTAACCAACAGAGATATGAGATGATGATGATGATGATGATGATGAT 2520  
 Db 2794 TCTGTGAGATTCAAATTAACCAACAGAGATATGAGATGATGATGATGATGATGATGATGAT 2853  
 Qy 2521 AGTGTGCTTCTCAACAG 2580  
 Db 2854 AGTGTGCTTCTCAACAG 2913  
 Qy 2581 CTGTGATTTCAACATCAAGCTTCTGGAACAGATCTCTTTAAATAGTCTTAAAGATTTGT 2640  
 Db 2914 CTGTGATTTCAACATCAAGCTTCTGGAACAGATCTCTTTAAATAGTCTTAAAGATTTGT 2973  
 Qy 2641 CTTCCTAATGAGACAACTTTTGAATCTGTGTTTGTGAAGAAATGTTGTGTGGCTTACA 2700  
 Db 2974 CTTCCTAATGAGACAACTTTTGAATCTGTGTTTGTGAAGAAATGTTGTGTGGCTTACA 3033  
 Qy 2701 CAGTTTAACTAGTGAAGCTGT 2760  
 Db 3034 CAGTTTAACTAGTGAAGCTGT 3093  
 Qy 2761 GCAGAGAGTCTTCTATCAGTTTAACTCAGCAAAATGTCAAACAACTAGGTATGAGAA 2820

Db 3094 GCAGAGTGTCTTCTATCAGTTATACCTCACCAAAATGTCAAAACAACTAGTATGAGAA 3153  
 Qy 2821 AATGAAAATTTCCACACATCAACAAAGAAATTAACAGTGTCTGTCTTCAATCTTTTG 2880  
 Db 3154 AATGAAAATTTCCACACATCAACAAAGAAATTAACAGTGTCTGTCTTCAATCTTTTG 3213  
 Qy 2881 ATGTTTCTAATCCGACTCTTAATTTTCATGTA 2913  
 Db 3214 ATGTTTCTAATCCGACTCTTAATTTTCATGTA 3246  
 RESULT 4  
 us-10-116-802-108  
 / Sequence 108, Application US/10116802  
 / Publication No. US20030065157A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Amy Lasek  
 / TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER  
 / FILE REFERENCE: PA-0045 US  
 / CURRENT APPLICATION NUMBER: US/10/116,802  
 / PRIOR FILING DATE: 2002-04-04  
 / PRIOR APPLICATION NUMBER: 60/281,593  
 / NUMBER OF SEQ ID NOS: 519  
 / SOFTWARE: PERL Program  
 / SEQ ID NO 108  
 / LENGTH: 3721  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / FEATURE:  
 / NAME/KEY: misc feature  
 / OTHER INFORMATION: Incyte ID No: 492750CB1  
 us-10-116-802-108  
 Query Match 99.8%; Score 2908.2; DB 13; Length 3721;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2910; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 ATGGGACCTGCAATCCGGGAGAGATGAGATTTTAAAGTTGAAATCTGTTGGTAA 60  
 Db 288 ATGGGACCTGCAATCCGGGAGAGATGAGATTTTAAAGTTGAAATCTGTTGGTAA 347  
 Qy 61 GGATCATTTGCTGTGTCTCAAGAGTGAATCCATTCACATGCTTTGGAAGTTGAATC 120  
 Db 348 GGATCATTTGCTGTGTCTCAAGAGTGAATCCATTCACATGCTTTGGAAGTTGAATC 407  
 Qy 121 AAAATGATATGAAG 180  
 Db 408 AAAATGATATGAAG 467  
 Qy 181 GTGAAAATATCATTCGCAATTTGAACATCCCTCTATCTGTGAGCTTTATTAATTTTGA 240  
 Db 468 GTGAAAATATCATTCGCAATTTGAACATCCCTCTATCTGTGAGCTTTATTAATTTTGA 527  
 Qy 241 GATGCAATTAATGTATCTGTGATTTAGAAATGTGCATATGAGAGAAATGAACAGTAT 300  
 Db 528 GATGCAATTAATGTATCTGTGATTTAGAAATGTGCATATGAGAGAAATGAACAGTAT 587  
 Qy 301 CTAAAGATTAAGAGTGAACCTTCTCAAGAAATGAAGCTGACATCTTACAGCAATC 360  
 Db 588 CTAAAGATTAAGAGTGAACCTTCTCAAGAAATGAAGCTGACATCTTACAGCAATC 647  
 Qy 361 ATCAGAGGATGTGTATCTTCAATCTCAATGATATACACCGGAGCTTCAATCAGATC 420  
 Db 648 ATCAGAGGATGTGTATCTTCAATCTCAATGATATACACCGGAGCTTCAATCAGATC 707  
 Qy 421 AACCTCTTACATCTGTATTAATGAACATCAAGATGCTGATTTTGGCTGGCACTCAA 480  
 Db 708 AACCTCTTACATCTGTATTAATGAACATCAAGATGCTGATTTTGGCTGGCACTCAA 767  
 Qy 481 CTGAAAATGCAATGAAAAGCACTATCATTAATGGAATCCTTAACTATTAATTAACCA 540  
 Db 768 CTGAAAATGCAATGAAAAGCACTATCATTAATGGAATCCTTAACTATTAATTAACCA 827

QY 541 GAAATGGCACTGGAAGTGCACATGGCTTGAATCTGAATGTTGGTCCCTGGGCTGATG 600  
DB 828 GAAATGGCACTGGAAGTGCACATGGCTTGAATCTGAATGTTGGTCCCTGGGCTGATG 887  
QY 601 TTTTATACATTAATTATCGGAGACCACTTTCAGACATGACACAGTCAAGAACATTA 660  
DB 888 TTTTATACATTAATTATCGGAGACCACTTTCAGACATGACACAGTCAAGAACATTA 947  
QY 661 AATAAGATGATGAGCAATTAAGAATGCCATCTTTTGTCAATAGAGGCCAAGAC 720  
DB 948 AATAAGATGATGAGCAATTAAGAATGCCATCTTTTGTCAATAGAGGCCAAGAC 1007  
QY 721 CTATATCACCAGTACTTGTGTGAATCCAGCAGATCTTAACTGTCTGTCAATG 780  
DB 1008 CTATATCACCAGTACTTGTGTGAATCCAGCAGATCTTAACTGTCTGTCAATG 1067  
QY 781 GACCATCTTTTATGTCGGAATCTTCAACAAAAAGTAAAGATTAGAGCTGGA 840  
DB 1068 GACCATCTTTTATGTCGGAATCTTCAACAAAAAGTAAAGATTAGAGCTGGA 1127  
QY 841 GACTCAATTGATAGTGGGCAATGCCAATTTCTACTGCAATTAAGCTTCTTCAAGTACC 900  
DB 1128 GACTCAATTGATAGTGGGCAATGCCAATTTCTACTGCAATTAAGCTTCTTCAAGTACC 1187  
QY 901 AGTATAGTGTAGTGTATTTTGAACAAAAGAGCTTTGTGTGAGAGCACTCCCAAT 960  
DB 1188 AGTATAGTGTAGTGTATTTTGAACAAAAGAGCTTTGTGTGAGAGCACTCCCAAT 1247  
QY 961 AAATAGCTGATATTTCCAAAGATTAAGTCAATGATTTTCTTCTTCAAGAGATGA 1020  
DB 1248 AAATAGCTGATATTTCCAAAGATTAAGTCAATGATTTTCTTCTTCAAGAGATGA 1307  
QY 1021 AACAGTTTTTATCTCACTGGGGAATCAAGAAACCAATATAGTGAAGGGAGAGTA 1080  
DB 1308 AACAGTTTTTATCTCACTGGGGAATCAAGAAACCAATATAGTGAAGGGAGAGTA 1367  
QY 1081 ATTCAAGATGCAAGAAAGGCCAATCTGCACTCTGTAGCTTATCTCTGAT 1140  
DB 1368 ATTCAAGATGCAAGAAAGGCCAATCTGCACTCTGTAGCTTATCTCTGAT 1427  
QY 1141 AGATCTGGCACTCTATATAGTCACTCAAGCAAAAAATATATCAATGAAAGATGAC 1200  
DB 1428 AGATCTGGCACTCTATATAGTCACTCAAGCAAAAAATATATCAATGAAAGATGAC 1487  
QY 1201 TCAGCAGAAATGCTTCACTGTCCAAAAGATCAGAGAGGTGAAAATGAAGAGGTAC 1260  
DB 1488 TCAGCAGAAATGCTTCACTGTCCAAAAGATCAGAGAGGTGAAAATGAAGAGGTAC 1547  
QY 1261 TCACCCAGACAAACATGCGCAATTTTAACTTCTTAAAGAAAAGACATCCAGTAC 1320  
DB 1548 TCACCCAGACAAACATGCGCAATTTTAACTTCTTAAAGAAAAGACATCCAGTAC 1607  
QY 1321 TCTGATCTTTTGAAGACCTGATTAACATCAAGACCTCCATCATCTTTGTCCAGGA 1380  
DB 1608 TCTGATCTTTTGAAGACCTGATTAACATCAAGACCTCCATCATCTTTGTCCAGGA 1667  
QY 1381 AAAAATCTTTTTCATTTGCAAGCCGCACTCAAGCTGAAAACGTAACAAGTGT 1440  
DB 1668 AAAAATCTTTTTCATTTGCAAGCCGCACTCAAGCTGAAAACGTAACAAGTGT 1727  
QY 1441 GGAATCTGCAAAATTAATGCTCAATTAAGAAAATCACTGAAATGACAGCACTCA 1500  
DB 1728 GGAATCTGCAAAATTAATGCTCAATTAAGAAAATCACTGAAATGACAGCACTCA 1787  
QY 1501 AACCGGAACTTCCAGGGCCATTCAGATTTGCAAGAGACATCAAAAAATGCTGACT 1560  
DB 1788 AACCGGAACTTCCAGGGCCATTCAGATTTGCAAGAGACATCAAAAAATGCTGACT 1847  
QY 1561 GATACAAAAGTCAAAAAGAACTGTGATGCTTCTGATATGCAATCTGTAAAAACAGCA 1620  
DB 1848 GATACAAAAGTCAAAAAGAACTGTGATGCTTCTGATATGCAATCTGTAAAAACAGCA 1907

QY 1621 AATACCATGAATAATATGACTGCACTTCAAGTAAACCTGAGATAATCCAAAGATGT 1680  
DB 1908 AATACCATGAATAATATGACTGCACTTCAAGTAAACCTGAGATAATCCAAAGATGT 1967  
QY 1681 GTTTTGGCTCAGATCTCTTTCTGAAACAGACAGACATGAGGGTATGAGCCCAATGG 1740  
DB 1968 GTTTTGGCTCAGATCTCTTTCTGAAACAGACAGACATGAGGGTATGAGCCCAATGG 2027  
QY 1741 GGTATCAAAATGATCAATTAAGAACATTAACATCCCTGGTGTCTCAGAGTTAAA 1800  
DB 2028 GGTATCAAAATGATCAATTAAGAACATTAACATCCCTGGTGTCTCAGAGTTAAA 2087  
QY 1801 CCAATCAGACAGAAAACCAAAAAGGCTGTGTGAGACATCTTGAATCAGAGAGGTGT 1860  
DB 2088 CCAATCAGACAGAAAACCAAAAAGGCTGTGTGAGACATCTTGAATCAGAGAGGTGT 2147  
QY 1861 GTGAGCTTGTAAAGAGATGATCTCAAGAAATATGTAAGAAAGTTCTTCAAGATCT 1920  
DB 2148 GTGAGCTTGTAAAGAGATGATCTCAAGAAATATGTAAGAAAGTTCTTCAAGATCT 2207  
QY 1921 AGTATGGAATATGATCTATTTATCCAAAATGAGGTGAGGTTTCTCTGCT 1980  
DB 2208 AGTATGGAATATGATCTATTTATCCAAAATGAGGTGAGGTTTCTCTGCT 2267  
QY 1981 GATAGACCAACCTCAGCTTCTGCAACATCAAGAGGTAGACCTTGAATTAACAGAA 2040  
DB 2268 GATAGACCAACCTCAGCTTCTGCAACATCAAGAGGTAGACCTTGAATTAACAGAA 2327  
QY 2041 AAATCTGGCGAAATATGATATGCTTCAAGTTTGTACAGTTGTAAAGATTAATCT 2100  
DB 2328 AAATCTGGCGAAATATGATATGCTTCAAGTTTGTACAGTTGTAAAGATTAATCT 2387  
QY 2101 CCCAAAATCACTATTTTCAAGATATGCTAAATGATTTTGAATGAAATCTCTGCT 2160  
DB 2388 CCCAAAATCACTATTTTCAAGATATGCTAAATGATTTTGAATGAAATCTCTGCT 2447  
QY 2161 GCTGATTTGAGGTTTGTATGATGAGGGTAAATATCACAAAACAGAAAGTTCAAT 2220  
DB 2448 GCTGATTTGAGGTTTGTATGATGAGGGTAAATATCACAAAACAGAAAGTTCAAT 2507  
QY 2221 CAGGTATGAAAGACAGGGAAGTCTTACCTTAAAGATGAAGTTAATAGC 2280  
DB 2508 CAGGTATGAAAGACAGGGAAGTCTTACCTTAAAGATGAAGTTAATAGC 2567  
QY 2281 TTGAAAAGAGATTAATAATGATATGAGACATGCTAAAGAGGTGATGTTGTTA 2340  
DB 2568 TTGAAAAGAGATTAATAATGATATGAGACATGCTAAAGAGGTGATGTTGTTA 2627  
QY 2341 GCACGTGAATCCATATTTCAAGAAAGAAAGAAATAGAGAGTCTCCCTTTCCCA 2400  
DB 2628 GCACGTGAATCCATATTTCAAGAAAGAAAGAAATAGAGAGTCTCCCTTTCCCA 2687  
QY 2401 ATATATCATGAGAAAGAAACCTGTAGTACTAGTTCACTTAAGGCTTATCACTCTCT 2460  
DB 2688 ATATATCATGAGAAAGAAACCTGTAGTACTAGTTCACTTAAGGCTTATCACTCTCT 2747  
QY 2461 TCTGTGATTTCAAAATTAACCAACAGAGATAGAGCATTTTCAACGAATGTCAGCAT 2520  
DB 2748 TCTGTGATTTCAAAATTAACCAACAGAGATAGAGCATTTTCAACGAATGTCAGCAT 2807  
QY 2521 AGTGTGCTTCTTCAACAGGCAACCAATCCTTAATCCCTTAATGTTGTTCAATGAAGA 2580  
DB 2808 AGTGTGCTTCTTCAACAGGCAACCAATCCTTAATCCCTTAATGTTGTTCAATGAAGA 2867  
QY 2581 CTGTGTCTTAAACATCAAGCTTCTGGAACAGACATCTCTTAAATGCTAAAGATGT 2640  
DB 2868 CTGTGTCTTAAACATCAAGCTTCTGGAACAGACATCTCTTAAATGCTAAAGATGT 2927  
QY 2641 CTTCCTTAAATCAAGCAAACTTTGAAATGCTTTTGTGAAAAAATGTTGGGCTACA 2700  
DB 2928 CTTCCTTAAATCAAGCAAACTTTGAAATGCTTTTGTGAAAAAATGTTGGGCTACA 2987  
QY 2701 CAGTTAACTAGTGAAGTGTGAGTTCAAGTTTAAATGATAGGCTCCAGTTGTTGTGAG 2760



Db 2388 CAGTAACTAGTGGAGCTGTGGCTTCAGTTTATGATGGGTCCAGTTGTGTGAG 3047  
 Qy 2761 GCAGAGAGTCTTCTATCAGTTATACCTCCAGCAATAGTCAAAACATAGTATGAGAA 2820  
 Db 3048 GCAGAGAGTCTTCTATCAGTTATACCTCCAGCAATAGTCAAAACATAGTATGAGAA 3107  
 Qy 2821 AATGAAAAATTCACGACTACATCAACAGAAATTAAGTGTGTCTTCATCCTTTTG 2880  
 Db 3108 AATGAAAAATTCACGACTACATCAACAGAAATTAAGTGTGTCTTCATCCTTTTG 2880  
 Qy 2881 AATGTTCTATCCGACTCCTTAATTTTCATTGA 2913  
 Db 3168 AATGTTCTATCCGACTCCTTAATTTTCATTGA 3200

RESULT 5  
 US-10-425-114-26254  
 ; Sequence 26254, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E.  
 ; APPLICANT: Tabaska, Jack E.  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(5313) B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 26254  
 ; LENGTH: 2836  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: LIB4119-112-B4\_FLI  
 ; US-10-425-114-26254

Query Match 95.7%; Score 2786.4; DB 13; Length 2836;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2787; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 126 GATAGTAAAGAACCATGTACAAAGCAGAAATGTGTACAGAGTCCAAATAGAGTGA 185  
 Db 1 GATAGTAAAGAACCATGTACAAAGCAGAAATGTGTACAGAGTCCAAATAGAGTGA 60  
 Qy 186 AATPACCTTGCATTAAGAACATCCTTCTATCTTGTGAGCTTTATATCTATTTGAAGATG 245  
 Db 61 AATPACCTTGCATTAAGAACATCCTTCTATCTTGTGAGCTTTATATCTATTTGAAGATG 120  
 Qy 246 CAATTATGTATCTGTATTAAGAAATGTGCCATATGAGAAATGAAACAGGTATCTAAA 305  
 Db 121 CAATTATGTATCTGTATTAAGAAATGTGCCATATGAGAAATGAAACAGGTATCTAAA 180  
 Qy 306 GAATGAGTGAACCCCTTCTCAGAAATGAAAGCTGACACTTCATGACCAAGATCATAC 365  
 Db 181 GAATGAGTGAACCCCTTCTCAGAAATGAAAGCTGACACTTCATGACCAAGATCATAC 240  
 Qy 366 AGGAGTGTGTATCTTCAATCTCATGTATCTACACGGGAGCTTCAACATTTCTAACCT 425  
 Db 241 AGGAGTGTGTATCTTCAATCTCATGTATCTACACGGGAGCTTCAACATTTCTAACCT 300  
 Qy 426 CTTACTGACTGTATATGAAACATCAAGATTGCTGATTTGGGCTGGCACTCAACTGAA 485  
 Db 301 CTTACTGACTGTATATGAAACATCAAGATTGCTGATTTGGGCTGGCACTCAACTGAA 360  
 Qy 486 AATGCAATGAAAGCACTATACCTTATGTGAACTCTTAATCAATTTCAAGAAAT 545  
 Db 361 AATGCAATGAAAGCACTATACCTTATGTGAACTCTTAATCAATTTCAAGAAAT 420

Qy 546 TGGCACTCGAAGTGCACATGCGCTTGAATCTGATGTTTGTCCCTGGGCTGTATGTTTAA 605  
 Db 421 TGGCACTCGAAGTGCACATGCGCTTGAATCTGATGTTTGTCCCTGGGCTGTATGTTTAA 480  
 Qy 606 TACATTACTTATCGGGAGACACCTTGCAGACTGACACAGTCAAGAAACATTTAAATA 665  
 Db 481 TACATTACTTATCGGGAGACACCTTGCAGACTGACACAGTCAAGAAACATTTAAATA 540  
 Qy 666 AGTAGATTTGGCAATTTATGAATGCCATCTTTTGTCAATATAGGCCCAAGACCTTAT 725  
 Db 541 AGTAGATTTGGCAATTTATGAATGCCATCTTTTGTCAATATAGGCCCAAGACCTTAT 600  
 Qy 726 TCACCAAGTACTCTGTATGAATTCACAGATGCTTAAGTCTGTCTTCAATTTGAGCA 785  
 Db 601 TCACCAAGTACTCTGTATGAATTCACAGATGCTTAAGTCTGTCTTCAATTTGAGCA 660  
 Qy 786 TCCCTTTATGTCGCCAAATTCCTTCAACAAAAGTAAAGATTAGGAACTGTGAGACTC 845  
 Db 661 TCCCTTTATGTCGCCAAATTCCTTCAACAAAAGTAAAGATTAGGAACTGTGAGACTC 720  
 Qy 846 AATGATATGAGGACATGCGCAATTTCTAGTCAATTACAGCTTCTTCAAGTACAGTAT 905  
 Db 721 AATGATATGAGGACATGCGCAATTTCTAGTCAATTACAGCTTCTTCAAGTACAGTAT 780  
 Qy 906 AAGTGTAGTTTATTTGACAAAGAACCTTTGATGTGAGGCACTCCCAATTTAAAT 965  
 Db 781 AAGTGTAGTTTATTTGACAAAGAACCTTTGATGTGAGGCACTCCCAATTTAAAT 840  
 Qy 966 GACTGTATTTCCAAAGAAATTAAGTTCACTGATTTTCTTCCAGGAGATGAAACAG 1025  
 Db 841 GACTGTATTTCCAAAGAAATTAAGTTCACTGATTTTCTTCCAGGAGATGAAACAG 900  
 Qy 1026 TTTTATATCTCAGTGGGGAATTCAGAAACCAAGTAAATGTGAGAGGGAGAGTAAATCA 1085  
 Db 901 TTTTATATCTCAGTGGGGAATTCAGAAACCAAGTAAATGTGAGAGGGAGAGTAAATCA 960  
 Qy 1086 AGATGCAAGAAAGGCCACATCTCTGATACCTTCTGATAGCTTATTCCTGTATGATC 1145  
 Db 961 AGATGCAAGAAAGGCCACATCTCTGATACCTTCTGATAGCTTATTCCTGTATGATC 1020  
 Qy 1146 TGGCACTTATATGTCAGTCTCAAGCAAAACATATACATGGAACGATGTCACTGAGC 1205  
 Db 1021 TGGCACTTATATGTCAGTCTCAAGCAAAACATATACATGGAACGATGTCACTGAGC 1080  
 Qy 1206 AGAATGCTTCACTGTCTCAAAAGATCAGAGGAGGGAATGAAGAGGATCTCAC 1265  
 Db 1081 AGAATGCTTCACTGTCTCAAAAGATCAGAGGAGGGAATGAAGAGGATCTCAC 1140  
 Qy 1266 CACAGACAAACATGCCAACTTTTAACTTCTTAAAGAAAGACATCCAGTATGTTCTGG 1325  
 Db 1141 CACAGACAAACATGCCAACTTTTAACTTCTTAAAGAAAGACATCCAGTATGTTCTGG 1200  
 Qy 1326 ATCTTTGAAAGACCTGTATTAACATCAAGACTCTCAATCATCTTTGTCCAGAAAAAC 1385  
 Db 1201 ATCTTTGAAAGACCTGTATTAACATCAAGACTCTCAATCATCTTTGTCCAGAAAAAC 1260  
 Qy 1386 TCCCTTTTCAATTTGCAACCCGACACTCAGACTGAAACCGTCAACAGGTTTGGGAA 1445  
 Db 1261 TCCCTTTTCAATTTGCAACCCGACACTCAGACTGAAACCGTCAACAGGTTTGGGAA 1320  
 Qy 1446 TCTGCAAAATATATGCTCATTTAAGAAAACTACTGATATGACAGATCAGCCCAACCG 1505  
 Db 1321 TCTGCAAAATATATGCTCATTTAAGAAAACTACTGATATGACAGATCAGCCCAACCG 1380  
 Qy 1506 GGACTTCAGGGGCATTCAGATTTTGAAGAGGACATCAAAAAAGCTGTGACTGATAC 1565  
 Db 1381 GGACTTCAGGGGCATTCAGATTTTGAAGAGGACATCAAAAAAGCTGTGACTGATAC 1440  
 Qy 1566 AAAAGTCAAAAAGAACTGTATGCTTCTGATATGACATCTGTATTAACAGCAAAATAC 1625  
 Db 1441 AAAAGTCAAAAAGAACTGTATGCTTCTGATATGACATCTGTATTAACAGCAAAATAC 1500  
 Qy 1626 CATGAAATATATGACTGACTTCAAGTAAACCTGATATATCCAAAGAAATGTGTGTTT 1685

```

Db 1501 CATGAATATATATGACCTGCACTTACAGTAACTGATATATCAACAAAGATGTGTTT 1560
Qy 1686 TGGCTCAGATCCTCTTTCTGGAACAGCAAGACTAGGGGTATGAGCCACCATGGGTTA 1745
Db 1561 TGGCTCAGATCCTCTTTCTGGAACAGCAAGACTAGGGGTATGAGCCACCATGGGTTA 1620
Qy 1746 TCAGATCGTACATTAAGAGCATTAATCTCGGTGTGTTGCTCAGAGGTTAAACCAT 1805
Db 1621 TCAGATCGTACATTAAGAGCATTAATCTCGGTGTGTTGCTCAGAGGTTAAACCAT 1680
Qy 1806 CAGACAAAACCAAAAAGGCTGTGTGAGCATCTTGAATTCAGAGAGGTGTGTGA 1865
Db 1681 CAGACAAAACCAAAAAGGCTGTGTGAGCATCTTGAATTCAGAGAGGTGTGTGA 1740
Qy 1866 GCTTGAAGAGATATCATCTCAAGAAATGTGAAGAAATTCCTCAATATCTAGTGA 1925
Db 1741 GCTTGAAGAGATATCATCTCAAGAAATGTGAAGAAATTCCTCAATATCTAGTGA 1800
Qy 1926 TGGAAATAGATCAGATTTATATCAATGGGTGAGAGGTTTCTCTGCTGATAG 1985
Db 1801 TGGAAATAGATCAGATTTATATCAATGGGTGAGAGGTTTCTCTGCTGATAG 1860
Qy 1986 ACCACCTCAGCTACTGACAAACATAGTAGTACAGCTTGAACAATTTACAGAAAAA 2045
Db 1861 ACCACCTCAGCTACTGACAAACATAGTAGTACAGCTTGAACAATTTACAGAAAAA 1920
Qy 2046 CTGGCGAAAAATCATATATGCTTCCAGGTTTGTACAGCTTTGAAGTCTAAATCTCC 2105
Db 1921 CTGGCGAAAAATCATATATGCTTCCAGGTTTGTACAGCTTTGAAGTCTAAATCTCC 1980
Qy 2106 AATCACTTATTTTACAGATATGCTAAATGCAATTTGATGAGAAATTCCTGCTGCTGA 2165
Db 1981 AATCACTTATTTTACAGATATGCTAAATGCAATTTGATGAGAAATTCCTGCTGCTGA 2040
Qy 2166 TTTTGAAGTTGGTTTATGATGAGGTTAAATAATACAAACAGAGATTTCAAGGT 2225
Db 2041 TTTTGAAGTTGGTTTATGATGAGGTTAAATAATACAAACAGAGATTTCAAGGT 2100
Qy 2226 GATTGAAGAGCAGGAGAGCTTTTACCTTTAAAGTGAAGTGAAGTTATAGCTTGA 2285
Db 2101 GATTGAAGAGCAGGAGAGCTTTTACCTTTAAAGTGAAGTGAAGTTATAGCTTGA 2160
Qy 2286 AAGAGGATTAATAATGTATATGAGCAATGCTAAAGGTCATGATTTGTTTACACT 2345
Db 2161 AAGAGGATTAATAATGTATATGAGCAATGCTAAAGGTCATGATTTGTTTACACT 2220
Qy 2346 GGAATTCATATTTTCAAGAGAGAGAAAGAACTAGAGAGTCTCCCTTTTCCCAATAT 2405
Db 2221 GGAATTCATATTTTCAAGAGAGAGAAAGAACTAGAGAGTCTCCCTTTTCCCAATAT 2280
Qy 2406 CATAGAGAGAAAACTGTAGTATCTAAGTTCACTTAAGGCTTTTCACTCTCTCTGT 2465
Db 2281 CATAGAGAGAAAACTGTAGTATCTAAGTTCACTTAAGGCTTTTCACTCTCTCTGT 2340
Qy 2466 GGAATTCATATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2525
Db 2341 GGAATTCATATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
Qy 2526 TGGCTTCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2585
Db 2401 TGGCTTCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460
Qy 2586 TCTTACACTACAGCTTCTGGAACAGACATCTCTTCTAATAGTCTAATAAGATGTCTCC 2645
Db 2461 TCTTACACTACAGCTTCTGGAACAGACATCTCTTCTAATAGTCTAATAAGATGTCTCC 2520
Qy 2646 TAAATACAGACAACTTTTGAATCTGTTTGTGAAGAAATGTGTGGCTACAGAT 2705
Db 2521 TAAATACAGACAACTTTTGAATCTGTTTGTGAAGAAATGTGTGGCTACAGAT 2580
Qy 2706 AACTAGTGAAGCTGTGTGGCTTCAAGTTATATGATGGGCTCCAGTGGTTGTGACAGCAG 2765

```

```

Db 2581 AACTAGTGAAGCTGTGTGGCTTCAAGTTATATGATGGGCTCCAGTTGGTTGACAGCAG 2640
Qy 2766 AGTGTCTTATCAGTTATATCTACCAACAAATGCTAAACAACTAGGTATGAGAAATGA 2825
Db 2641 AGTGTCTTATCAGTTATATCTACCAACAAATGCTAAACAACTAGGTATGAGAAATGA 2700
Qy 2826 AAAATTCACAGACTATCAATCAACAGAAATTTACAGTGTCTGTCTTCATCTTTTGAAT 2885
Db 2701 AAAATTCACAGACTATCAATCAACAGAAATTTACAGTGTCTGTCTTCATCTTTTGAAT 2760
Qy 2886 TTCTAATCCGACTCTTAATTTTCAATGA 2913
Db 2761 TTCTAATCCGACTCTTAATTTTCAATGA 2788

RESULT 6
US-10-062-674-1868
; Sequence 1868, Application US/10062674
; Publication No. US2004000559A1
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.; Kager, Matthew R.
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS
; FILE REFERENCE: PA-0026-1 CIP
; CURRENT APPLICATION NUMBER: US/10/062,674
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: US 09/625,102
; NUMBER OF SEQ ID NOS: 2217
; SOFTWARE: PERL Program
; SEQ ID NO 1868
; LENGTH: 3924
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US2004000559A1 332518.2
US-10-062-674-1868

Query Match 92.2%; Score 2686.8; DB 16; Length 3924;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 2886; Conservative 0; Mismatches 27; Indels 23; Gaps 16;

Qy 1 ATGGCAGCTGCATCGGGGAGAGATCGAGATTTTAAAGTTGAATCTGCTGGTAA 60
Db 283 ATGGCAGCTGCATCGGGGAGAGATCGAGATTTTAAAGTTGAATCTGCTGGTAA 348
Qy 61 GGATCATTTGCTGTGTCTACAGAGCTGATTCATTCACATGCTGTTGGAGTTGCAAT 119
Db 349 GGATCATTTGCTGTGTCTACAGAGCTGATTCATTCACATGCTGTTGGAGTTGCAAT 408
Qy 120 CAAATGATATGATTAAGAAAGCCATGTACAAACAGAGATGTATCAGAGTCCAAATGA 179
Db 409 CAAATGATATGATTAAGAAAGCCATGTATCAAAACAGAGATGTATCAGAGTCCAAATGA 468
Qy 180 GGTGAATATACATTCGCAATTTGAACATCTCTTATCTTGAAGTTTATATATTTGA 239
Db 469 GGTGAATATACATTCGCAATTTGAACATCTCTTATCTTGAAGTTTATATATTTGA 528
Qy 240 AGATAGCAATTTATGTATCTGTATTTAGAAATGTGCCATTAATGAGAAATGAACAGTA 299
Db 529 AGATAGCAATTTATGTATCTGTATTTAGAAATGTGCCATTAATGAGAAATGAACAGTA 588
Qy 300 TCTTAAGAAATAGAGGAAACCTTTCTCAGAAATAGAGTGCACATTCATATGCAACAGT 359
Db 589 TCTTAAGAAATAGAGGAAACCTTTCTCAGAAATAGAGTGCACATTCATATGCAACAGT 648
Qy 360 CATCACAGGATGTGTATCTTCAATTCATGTATGATATCAACCGGAGCTCACAATTT 419
Db 649 CATCACAGGATGTGTATCTTCAATTCATGTATGATATCAACCGGAGCTCACAATTT 708
Qy 420 TAACTCTTACTGACTCGTAAATATGAACATCAAGATTTGCTGATTTTGGCTGGCACTGA 479
Db 709 TAACTCTTACTGACTCGTAAATATGAACATCAAGATTTGCTGATTTTGGCTGGCACTGA 768

```



QY 480 ACTGAAATGCGCAGTAAAGACCTATACATTATGTGGAATCTCTAACTACATTTCACC 539  
DB 769 ACTGAAAAATCGCAGATGAAAAAGACCTATACATTATGTGGAATCTCTAACTACATTTCACC 828  
QY 540 AGAATATGCGCACTCGAAGTGCACATGGCCTTGAAATCTGATGTGGTCCCGGGCTGAT 599  
DB 829 AAAAAATGCGCACTCGAAGTGCACATGGCCTTGAAATCTGATGTGGTCCCGGGCTGAT 888  
QY 600 GTTTATATACATTACTATTCGGGAGACCACTCTGACACTGACACAGTCAAGAACACATT 659  
DB 889 GTTTATATACATTACTATTCGGGAGACCACTCTGACACTGACACAGTCAAGAACACATT 948  
QY 660 AAAATAAGTATGTGGCAGATTATGAATGCACTCTTTTGTCAATAGAGGCGCAAGGA 719  
DB 949 AAAATAAGTATGTGGCAGATTATGAATGCACTCTTTTGTCAATAGAGGCGCAAGGA 1008  
QY 720 CTTATTCACAGATTACTCTGTGTAAGAAATCCGACGATCGTTTAAAGTGTCTTCAGTATT 779  
DB 1009 CTTATTCACAGATTACTCTGTGTAAGAAATCCGACGATCGTTTAAAGTGTCTTCAGTATT 1068  
QY 780 GGACCATCTCTTTTATGTCCGAAATCTTCACAAAAAGTAAAGATTAGAACTGGA 839  
DB 1069 GGACCATCTCTTTTATGTCCGAAATCTTCACAAAAAGTAAAGATTAGAACTGGA 1128  
QY 840 AGACTCAATTGATATGTGGGCAATGCGCAAAATTTCTACTGCAATTACAGCTTCTTCAGTAC 899  
DB 1129 AGACTCAATTGATATGTGGGCAATGCGCAAAATTTCTACTGCAATTACAGCTTCTTCAGTAC 1188  
QY 900 CAGTATAGTGTAGTATTATTTTGAACAAAAGACTTTTGAATGTGTCAGCACTCCCAA 959  
DB 1189 CAGTATAGTGTAGTATTATTTTGAACAAAAGACTTTTGAATGTGTCAGCACTCCCAA 1248  
QY 960 TAAAAATACCTGTATTTCCAAAAGATAAAAGTTCACTGATTTTCTTCTTCAGAGATG 1019  
DB 1249 TAAAAATACCTGTATTTCCAAAAGATAAAAGTTCACTGATTTTCTTCTTCAGAGATG 1308  
QY 1020 AAACAGTTTATTAATCAAGTGGGAAATCAAGAAACAGTAATATGTGAAAGGGAAGT 1079  
DB 1309 AAACAGTTTATTAATCAAGTGGGAAATCAAGAAACAGTAATATGTGAAAGGGAAGT 1368  
QY 1080 AATTCAGATGCAAGAAAGGCGCAGATTCGATACCTTCGTAAGCTTATTCCTGGA 1139  
DB 1369 AATTCAGATGCAAGAAAGGCGCAGATTCGATACCTTCGTAAGCTTATTCCTGGA 1428  
QY 1140 TAGATCTGGCACTTCAATATGCTCAAGCTCAAGCAAAACATATACATGGAACGATGCA 1199  
DB 1429 TAGATCTGGCACTTCAATATGCTCAAGCTCAAGCAAAACATATACATGGAACGATGCA 1488  
QY 1200 CTCAGCAGAATGCTTTCAGTGTCCAAAGATCAGAGAGGTGAAAAAGAGAGGTA 1259  
DB 1489 CTCAGCAGAATGCTTTCAGTGTCCAAAGATCAGAGAGGTGAAAAAGAGAGGTA 1548  
QY 1260 CTCACCCACAGACAAATGCGCAACTTTTAACTTTTAAAGAAAGACATCCAGTAG 1319  
DB 1549 CTCACCCACAGACAAATGCGCAACTTTTAACTTTTAAAGAAAGACATCCAGTAG 1608  
QY 1320 TTCTGATCTTTTGAAGACTGATACATCAAGCACTCCCAAT-CATCTTGTCCAG 1378  
DB 1609 TTCTGATCTTTTGAAGACTGATACATCAAGCACTCCCAAT-CATCTTGTCCAG 1668  
QY 1379 GAAAAATCTCTTTTCAATTTGACAGCCGACACTGAGAAACGTAACAAGTGT 1438  
DB 1669 GAAAAATCTCTTTTCAATTTGACAGCCGACACTGAGAAACGTAACAAGTGT 1728  
QY 1439 TTGGGAATCTGCAAAATAATGCTCATTTAAGAAAACTAAGTAATAGACATCAAGCC 1498  
DB 1729 TTGGGAATCTGCAAAATAATGCTCATTTAAGAAAACTAAGTAATAGACATCAAGCC 1788  
QY 1499 CAAACCGGGAATTCAGAGGCGCATCCAGATTTGCAAGAGACACATCAAAAAATGCCGGA 1558  
DB 1789 CAAACCGGGAATTCAGAGGCGCATCCAGATTTGCAAGAGACACATCAAAAAATGCCGGA 1848

QY 1559 CTGATACAAAAGTCAAAAAGACTGATGCTTGATATAGCACTTCTGTAACAGC 1618  
DB 1849 CTGATACAAAAGTCAAAAAGACTGATGCTTGATATAGCACTTCTGTAACAGC 1908  
QY 1619 AAAATACATGAATATATGACTGCACTTCACAGTAACCTGAGATTAATCCAAAGAAAT 1678  
DB 1909 AAAATACATGAATATATGACTGCACTTCACAGTAACCTGAGATTAATCCAAAGAAAT 1968  
QY 1679 GTGTTTATGCTCAGATCTCTCTTCGAAACAGACAGACTATGGGGATAGAGCCACAT 1738  
DB 1969 GTGTTTATGCTCAGATCTCTCTTCGAAACAGACAGACTATGGGGATAGAGCCACAT 2028  
QY 1739 GGGGTATCAGATCGTACATTATAGAGCACTTACCTCCGTGGTCTCAAGGTTAA 1798  
DB 2029 GGGGTATCAGATCGTACATTATAGAGCACTTACCTCCGTGGTCTCAAGGTTAA 2088  
QY 1799 AACCAATCAGACAGAAAAACCAAAAAAGCTGTGGTGAAGATACCTTGANTCAGAGAGGTG 1858  
DB 2089 AACCAATCAGACAGAAAAACCAAAAAAGCTGTGGTGAAGATACCTTGANTCAGAGAGGTG 2148  
QY 1859 GTGTGAAGCTGTAAAGAGATGCACTCAAGAATATGTGAAGAAAGTTCTTGATAT 1918  
DB 2149 GTGTGAAGCTGTAAAGAGATGCACTCAAGAATATGTGAAGAAAGTTCTTGATAT 2208  
QY 1919 CTAGATAGCAAAATACATCACTATTTATTCCAAATGGT-GGTAGAGTTTCTCTT 1977  
DB 2209 CTAGATAGCAAAATACATCACTATTTATTCCAAATGGTGGGTAGAGTTTCTCTT 2268  
QY 1978 GCT-GATAGACCACTTCACCTAC-TGACAAATCAATAGTATGACGTTTGAAGATTA 2035  
DB 2269 GCTGATAGCACTTCACCTTCACCTACCTTACCAATCAATAGTATGACGTTTGAAGATTA 2328  
QY 2036 CAGAAAAATATGCGGCAAAATATCAATATGCTTCAGGTTGTACAGCTTGAAGATTA 2095  
DB 2329 CAGAAAAATATGCGGCAAAATATCAATATGCTTCAGGTTGTACAGCTTGAAGATTA 2388  
QY 2096 AATCTCCAAAATCACTTATTTTACAAAGATATGCAATGCAATTTTGTATGAGAAATCTC 2155  
DB 2389 AATCTCCAAAATCACTTATTTTACAAAGATATGCAATGCAATTTTGTATGAGAAATCTC 2448  
QY 2156 CTGGTCTGATTTTGAAGTTGTGTTTATGATGGGTAAAAATACCAAAAAGAGATT 2215  
DB 2449 CTGGTCTGATTTTGAAGTTGTGTTTATGATGGGTAAAAATACCAAAAAGAGATT 2508  
QY 2216 TCATTCAGATGATGAAAAAGACAGGAAGCTTACCTTTAAAGTGAAGTGAAGTTA 2275  
DB 2509 TCATTCAGATGATGAAAAAGACAGGAAGCTTACCTTTAAAGTGAAGTGAAGTTA 2568  
QY 2276 ATAGCTTGAAGAGAGATTAATAATGTAT-ATGACCATGCTAAT-GAGGTCATCGTAT 2333  
DB 2569 ATAGCTTGAAGAGAGATTAATAATGTATATAGACCATGCTAATGAGGTCATCGTAT 2628  
QY 2334 TTGTTTAGACCTGGGAATCCATATTTTCAAGAGAGAAAGAAACCTAGAGTGTCTCCTT 2393  
DB 2629 TTGTTTAGACCTGGGAATCCATATTTTCAAGAGAGAAAGAAACCTAGAGTGTCTCCTT 2688  
QY 2394 TTTCCTCAATATCATAGAGAAAGAACTGGT-AGTACTAGTTCACTTAAGGCTTATAC 2452  
DB 2689 TTTCCTCAATATCATAGAGAAAGAACTGGTAGTACTAGTTCACTTAAGGCTTATAC 2748  
QY 2453 CTCCTCTCTC-1GTGATTCAAATTAACCAACGAGAGATAGAGACTTTTCAACAGATG 2511  
DB 2749 CTCCTCTCTC-1GTGATTCAAATTAACCAACGAGAGATAGAGACTTTTCAACAGATG 2808  
QY 2512 GTCATGATAGTGTGCTTCTCCAAACAGAGCAACAACTTAAATCCCTCTATAGTTACA 2571  
DB 2809 GTCATGATAGTGTGCTTCTCCAAACAGAGCAACAACTTAAATCCCTCTATAGTTACA 2868  
QY 2572 AATGAAGACTTGTGCTTACAACTACAGCTTCT-GGAACAGACATCTTCTAATAGTCT 2630  
DB 2869 AATGAAGACTTGTGCTTACAACTACAGCTTCTTGGGACACAGACATCTTCTAATAGTCT 2928  
QY 2631 AAAAGATGTCTTCTCTAAA---TGACGACAACTTTGAATC---TGTTTTGTGAAAAA 2684

Db 2929 AAAAGATGTCCTTCATATATACGACCAACCTTTGAAATCTGTTTTGTGAAAAAAT 2988  
Qy 2685 TGTGGTGGGCTACACACCTTAACTAGTGTGTGGGTTCACTTAATGATG 2740  
Db 2989 GTTGGTGGGCTACACACCTTAACTAGTGTGTGGGTTCACTTAATGATG 3048  
Qy 2741 GGTCCAGTTGGTGTGACGAGGAG-GAGTGTCTTATC-AGTTATACCTCACCAATGG 2798  
Db 3049 GGTCCAGTTGGTGTGACGAGGAGTGTCTTATACAGTTATACCTCACCAATGG 3108  
Qy 2799 TCACA-CAACTAGTATGAGAAATGAAATTAACAGACTACATCAACAGAAATAC 2857  
Db 3109 TCACAACCACTAGTATGAGAAATGAAATTAACAGACTACATCAACAGAAATAC 3168  
Qy 2858 AGTGTGTCTTCACTCTTTGATGATGTTTCTAATCGACTCTTAATTTTCAATGA 2913  
Db 3169 AGTGTGTCTTCACTCTTTGATGATGTTTCTAATCGACTCTTAATTTTCAATGA 3224

## RESULT 7

US-09-918-995-21560  
; Sequence 21560, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hybreg, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 21560  
; LENGTH: 484  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(484)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-21560

Query Match 14.6%; Score 426; DB 10; Length 484;

Best Local Similarity 100.0%; Pred. No. 5.2e-101;  
Matches 426; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2026 GACAATTTACGAAATATATCTGCGAATATCAATGCTCCAGTTTGTACAGCTT 2085  
Db 59 GACAATTTACGAAATATATCTGCGAATATCAATGCTCCAGTTTGTACAGCTT 118  
Qy 2086 GTAAAGATCTAAATCTCCAAATCACTTAATTTTACAAGATATCTAAATGATTTGATG 2145  
Db 119 GTAAAGATCTAAATCTCCAAATCACTTAATTTTACAAGATATCTAAATGATTTGATG 178  
Qy 2146 GAGAATTTCTGCTGCTGCTGATTTTGAAGTTTGTATGATGGGTAAATATACAAAA 2205  
Db 179 GAGAATTTCTGCTGCTGCTGATTTTGAAGTTTGTATGATGGGTAAATATACAAAA 238  
Qy 2206 AACAAGATTTCACTTACAGATGATTTGAAAGACAGGAAAGCTTACCTTTAAAGATGAA 2265  
Db 239 AACAAGATTTCACTTACAGATGATTTGAAAGACAGGAAAGCTTACCTTTAAAGATGAA 298  
Qy 2266 AGTGAAGTTAATAGCTTGAAGAGAGATTAATATATGACATGCTAATGAGGT 2325  
Db 299 AGTGAAGTTAATAGCTTGAAGAGAGATTAATATATGACATGCTAATGAGGT 358  
Qy 2326 CATGTAATTTGTTTACACTGGAATCCATTAATTTCAAGAGAGAAAGAAATAGAGAGT 2385  
Db 359 CATGTAATTTGTTTACACTGGAATCCATTAATTTCAAGAGAGAAAGAAATAGAGAGT 418

Qy 2386 GCTCCCTTTTCCCATATATATAGAGAAAACTGTAGTACTAGTTCACTTAAGGCC 2445  
Db 419 GCTCCCTTTTCCCATATATATAGAGAAAACTGTAGTACTAGTTCACTTAAGGCC 478  
Qy 2446 TTATCA 2451  
Db 479 TTATCA 484

## RESULT 8

US-09-878-178-96  
; Sequence 96, Application US/09878178  
; Patent No. US20020177552A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.527  
; CURRENT APPLICATION NUMBER: US/09/878,178  
; CURRENT FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 2237  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 96  
; LENGTH: 328  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-878-178-96

Query Match 11.3%; Score 328; DB 9; Length 328;

Best Local Similarity 100.0%; Pred. No. 2.2e-75;  
Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1428 ACAACAGTGTGTTGGATCTGCAATTAATGCTCTATTAAAGAAAACTAGTAATATGA 1487  
Db 1 ACAACAGTGTGTTGGATCTGCAATTAATGCTCTATTAAAGAAAACTAGTAATATGA 60  
Qy 1488 CAGCATCAGCCCAACCGGACTTCCAGGCGCATCCAGATTGCGAAGACATCAAA 1547  
Db 61 CAGCATCAGCCCAACCGGACTTCCAGGCGCATCCAGATTGCGAAGACATCAAA 120  
Qy 1548 AATGCTGAGTGTATCAAAAGTCAAAAAGACTGTGATGCTTGTATATGACATTC 1607  
Db 121 AATGCTGAGTGTATCAAAAGTCAAAAAGACTGTGATGCTTGTATATGACATTC 180  
Qy 1608 TGTAAACAGCAAAATATACATGAATATATGACTGCACTTCAAGTAACCTGAGTAAT 1667  
Db 181 TGTAAACAGCAAAATATACATGAATATATGACTGCACTTCAAGTAACCTGAGTAAT 240  
Qy 1668 CCAACAAAGATGTTTGTGCTCAGATCTCTTCTGAACAGAGAACTAGAGGATAT 1727  
Db 241 CCAACAAAGATGTTTGTGCTCAGATCTCTTCTGAACAGAGAACTAGAGGATAT 300  
Qy 1728 GGAGCCACCATGGGTTATCAGAAATGCT 1755  
Db 301 GGAGCCACCATGGGTTATCAGAAATGCT 328

## RESULT 9

US-10-046-935-96  
; Sequence 96, Application US/10046935  
; Publication No. US20020156011A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Wang, Aijun  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.527C1  
; CURRENT APPLICATION NUMBER: US/10/046,935

/ CURRENT FILING DATE: 2002-01-15  
 / NUMBER OF SEQ ID NOS: 2239  
 / SOFTWARE: FastSeq for Windows Version 4.0  
 / SEQ ID NO 96  
 / LENGTH: 328  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 US-10-046-935-96

Query Match 11.3%; Score 328; DB 14; Length 328;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-75;  
 Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1428 ACAACAGTGGTTGGGAATCTGCAAAATATATGCTCATTTAAGAAAACTACTGAATATGA 1487  
 DB 1 ACAACAGTGGTTGGGAATCTGCAAAATATATGCTCATTTAAGAAAACTACTGAATATGA 60  
 QY 1488 CAGCATCAGCCCAACCCGGGACTTCAGGGCCATCCAGATTGGAGAGACATCAAA 1547  
 DB 61 CAGCATCAGCCCAACCCGGGACTTCAGGGCCATCCAGATTGGAGAGACATCAAA 120  
 QY 1548 AATGCTGAGCTGATACAAAGTCAAAAAGACTGAGCTTCGATATGCACTTC 1607  
 DB 121 AATGCTGAGCTGATACAAAGTCAAAAAGACTGAGCTTCGATATGCACTTC 180  
 QY 1608 TGTAAAACAGCAAAATACCATGAATATATGACTGCACTTCAAGTAACTGAGATTAAT 1667  
 DB 181 TGTAAAACAGCAAAATACCATGAATATATGACTGCACTTCAAGTAACTGAGATTAAT 240  
 QY 1668 CCAACAAGATGTTGTTGGCTCAGATCCTCTTTCGAAACAGACGACATGAGGATAT 1727  
 DB 241 CCAACAAGATGTTGTTGGCTCAGATCCTCTTTCGAAACAGACGACATGAGGATAT 300  
 QY 1728 GAGGCCACCATGGGGTTATCAGATCGT 1755  
 DB 301 GAGGCCACCATGGGGTTATCAGATCGT 328

RESULT 10  
 US-10-146-502-96  
 / Sequence 96, Application US/10146502  
 / Publication No. US20030069180A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Jiang, Yujin  
 / APPLICANT: Harlocker, Susan L.  
 / APPLICANT: Secretist, Heather  
 / APPLICANT: Wang, Aijun  
 / APPLICANT: Stolk, John A.  
 / TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
 / FILE REFERENCE: 210121.527C2  
 / CURRENT APPLICATION NUMBER: US/10/146,502  
 / CURRENT FILING DATE: 2002-05-14  
 / NUMBER OF SEQ ID NOS: 2241  
 / SOFTWARE: FastSeq for Windows Version 4.0  
 / SEQ ID NO 96  
 / LENGTH: 328  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 US-10-146-502-96

Query Match 11.3%; Score 328; DB 15; Length 328;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-75;  
 Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1428 ACAACAGTGGTTGGGAATCTGCAAAATATATGCTCATTTAAGAAAACTACTGAATATGA 1487  
 DB 1 ACAACAGTGGTTGGGAATCTGCAAAATATATGCTCATTTAAGAAAACTACTGAATATGA 60  
 QY 1488 CAGCATCAGCCCAACCCGGGACTTCAGGGCCATCCAGATTGGAGAGACATCAAA 1547  
 DB 61 CAGCATCAGCCCAACCCGGGACTTCAGGGCCATCCAGATTGGAGAGACATCAAA 120

QY 1548 AATGCTGAGCTGATACAAAGTCAAAAAGACTGATGCTTCTGATATGCACTTC 1607  
 DB 121 AATGCTGAGCTGATACAAAGTCAAAAAGACTGATGCTTCTGATATGCACTTC 180  
 QY 1608 TGTAAAACAGCAAAATACCATGAATATATGACTGCACTTCAAGTAACTGAGATTAAT 1667  
 DB 181 TGTAAAACAGCAAAATACCATGAATATATGACTGCACTTCAAGTAACTGAGATTAAT 240  
 QY 1668 CCAACAAGATGTTGTTGGCTCAGATCCTCTTTCGAAACAGACGACATGAGGATAT 1727  
 DB 241 CCAACAAGATGTTGTTGGCTCAGATCCTCTTTCGAAACAGACGACATGAGGATAT 300  
 QY 1728 GAGGCCACCATGGGGTTATCAGATCGT 1755  
 DB 301 GAGGCCACCATGGGGTTATCAGATCGT 328

RESULT 11  
 US-10-062-674-914  
 / Sequence 914, Application US/10062674  
 / Publication No. US20040005559A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.  
 / TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS  
 / FILE REFERENCE: PA-0026-1 CIP  
 / CURRENT APPLICATION NUMBER: US/10/062,674  
 / CURRENT FILING DATE: 2002-01-30  
 / PRIOR APPLICATION NUMBER: US 09/625,102  
 / NUMBER OF SEQ ID NOS: 2217  
 / SOFTWARE: PERL Program  
 / SEQ ID NO 914  
 / LENGTH: 445  
 / TYPE: DNA  
 / ORGANISM: Mus musculus  
 / FEATURE:  
 / NAME/KEY: misc feature  
 / OTHER INFORMATION: Genbank ID No. US20040005559A1 g1919634  
 US-10-062-674-914

Query Match 9.5%; Score 276; DB 16; Length 445;  
 Best Local Similarity 81.2%; Pred. No. 1.3e-61;  
 Matches 359; Conservative 0; Mismatches 75; Indels 8; Gaps 3;

QY 1576 AAGAATCTGATGCTTCTGATATGCAATCTCTGTAACAGCAAAATACCATGAATAT 1635  
 DB 12 AAGAATCTGATGCTTCTGATATGCAATCTCTGTAACAGCAAAATACCATGAATAT 70  
 QY 1636 ATGACTGCACTTCAAGTAACTGAGATTAATCAACAAGATGTTGTTGGCTCAGAT 1695  
 DB 71 ATGACTGCACTTCAAGTAACTGAGATTAATCAACAAGATGTTGTTGGCTCAGAT 124  
 QY 1696 CTTCTTTCGAAACAGCAAGACTAGGGATATGAGGCCACATGAGGTTATCAGATCGT 1755  
 DB 125 CTTCTTTCGAAACAGCAAGACTAGGGATATGAGGCCACATGAGGTTATCAGATCGT 184  
 QY 1756 ACATTAAAGACATTACATCTCCGTTGTTGCTCAAGGTTAAACCATCAGAGAAA 1815  
 DB 185 ACATTAAAGACATTACATCTCCGTTGTTGCTCAAGGTTAAACCATCAGAGAAA 244  
 QY 1816 ACCAAAAAGGCTGTGTGAGCATTAATTCAGAGAGGTGTGTGAGGCTTGTAAAG 1875  
 DB 245 ACCAAAAAGGCTGTGTGAGCATTAATTCAGAGAGGTGTGTGAGGCTTGTAAAG 304  
 QY 1876 GAGTATGATCTCAAGATATATGTAAGAGTTCTTCGATATCTAGATGAGAAATGCG 1935  
 DB 305 GAGTATGATCTCAAGATATATGTAAGAGTTCTTCGATATCTAGATGAGAAATGCG 363  
 QY 1936 ATCACTATTATTAATCAAAATGTTGAGGTTTCTCTTTCGATATGAGCAACCTCA 1995  
 DB 364 ATCACTATTATTAATCAAAATGTTGAGGTTTCTCTTTCGATATGAGCAACCTCA 423  
 QY 1996 CCTACTGACAAACATCAGTAGT 2017

Db 424 CCGACTGACGACGACTGAGT 445

## RESULT 12

US-09-918-995-24717  
; Sequence 24717, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: HySeq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCES: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24717  
; LENGTH: 491  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)\_(491)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-24717

Query Match 8.6%; Score 250; DB 10; Length 491;  
Best Local Similarity 100.0%; Pred. No. 9e-55;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGACCTGCATCGGGGAGAAAGATCGAGATTTTAAAGTTGAAATCTGTTGTA 60  
DB 241 ATGGCGACCTGCATCGGGGAGAAAGATCGAGATTTTAAAGTTGAAATCTGTTGTA 300  
QY 61 GGATCATTTGCTGCTGTCTACAGAGCTGATCCATTCACATGTTGGAAGTTGCAATC 120  
DB 301 GGATCATTTGCTGCTGTCTACAGAGCTGATCCATTCACATGTTGGAAGTTGCAATC 360  
QY 121 AAAATGATGATTAAGAAAGCATGTACAAAGCAGATGTTAGACAGAGTCCAAAATGAG 180  
DB 361 AAAATGATGATTAAGAAAGCATGTACAAAGCAGATGTTAGACAGAGTCCAAAATGAG 420  
QY 181 GTGAAATATCATTTGCCAATTGAAACATCTTCTATCTTGAGCTTTATTAATTTGAA 240  
DB 421 GTGAAATATCATTTGCCAATTGAAACATCTTCTATCTTGAGCTTTATTAATTTGAA 480  
QY 241 GATAGCAATT 250  
DB 481 GATAGCAATT 490

## RESULT 13

US-10-152-319A-292/c  
; Sequence 292, Application US/10152319A  
; Publication No. US20040072160A1  
; GENERAL INFORMATION:  
; APPLICANT: Mendrick, Donna  
; APPLICANT: Porter, Mark  
; APPLICANT: Johnson, Kory  
; APPLICANT: Higgs, Brandon  
; APPLICANT: Castle, Arthur  
; APPLICANT: Blashoff, Michael  
; TITLE OF INVENTION: Molecular Toxicology Modeling  
; FILE REFERENCES: 44921-5089-US  
; CURRENT APPLICATION NUMBER: US/10/152,319A  
; PRIOR FILING DATE: 2002-05-22  
; PRIOR APPLICATION NUMBER: US 60/292,335  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/297,523  
; PRIOR FILING DATE: 2001-06-13

; PRIOR APPLICATION NUMBER: US 60/298,925  
; PRIOR FILING DATE: 2001-06-19  
; PRIOR APPLICATION NUMBER: US 60/303,810  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/303,807  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/303,808  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/315,047  
; PRIOR FILING DATE: 2001-08-28  
; PRIOR APPLICATION NUMBER: US 60/324,928  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US 60/330,867  
; PRIOR FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: US 60/330,462  
; PRIOR FILING DATE: 2001-10-22  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 2221  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 292  
; LENGTH: 302  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. AA894059  
US-10-152-319A-292

Query Match 7.4%; Score 216.8; DB 12; Length 302;  
Best Local Similarity 91.3%; Pred. No. 3.4e-46;  
Matches 230; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 2662 TTGAATCTGTTTGTGAAATGTTGTTGGCTACACAGTTACTAGTGAGCTG 2721  
DB 302 TTGAATCTGTTTGTGAAATGTTGTTGGCTACACAGTTACTAGTGAGCTG 243  
QY 2722 TGGGTTCACTTAATGATGGGTCCTCAGTTGTTGTCAGGCAAGATCTTCACT 2781  
DB 242 TGGGTTCACTTAATGATGGGTCCTCAGTTGTTGTCAGGCAAGATCTTCACT 183  
QY 2782 TATACCTCACCAATGTCGAAACACTAGGTATGGAATAATTAACAGCTAC 2841  
DB 182 TATACCTCACCAATGTCGAAACACTAGGTATGGAATAATTAACAGCTAC 123  
QY 2842 ATCAACAGAAATTAAGTGTCTTCATCTTGAATGTTTCTATCCAGCTCCT 2901  
DB 122 ATCAACAGAAATTAAGTGTCTTCATCTTGAATGTTTCTATCCAGCTCCT 63  
QY 2902 AATTTTCATTGA 2913  
DB 62 AGTTTTCAGTGA 51

## RESULT 14

US-10-305-720-649  
; Sequence 649, Application US/10305720  
; Publication No. US20040010136A1  
; GENERAL INFORMATION:  
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.  
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression  
; FILE REFERENCES: PA-0002-1 CON  
; CURRENT APPLICATION NUMBER: US/10/305,720  
; PRIOR FILING DATE: 2002-11-26  
; PRIOR APPLICATION NUMBER: 09/016,434  
; PRIOR FILING DATE: 1998-01-30  
; NUMBER OF SEQ ID NOS: 1490  
; SOFTWARE: PERL Program  
; SEQ ID NO 649  
; LENGTH: 381  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040010136A1 259999

Search completed: September 24, 2004, 20:47:13  
Job time : 1312 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 20:57:40 ; Search time 15 Seconds

(without alignments)  
3367.203 Million cell updates/sec

Title: US-10-026-021-2

Perfect score: 5078

Sequence: 1 MATCGEKEDPKVGNLGLK.....KLGCSLLMFSPNPNPH 970

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	599.5	11.8	576	1	POLO_DROME
2	593	11.7	631	1	POLO_MOUSE
3	590.5	11.6	615	1	POLO_MOUSE
4	589.5	11.6	646	1	POLO_MOUSE
5	561.5	11.1	682	1	POLO_MOUSE
6	560.5	11.0	682	1	POLO_MOUSE
7	554.5	11.0	685	1	POLO_MOUSE
8	548	10.8	408	1	POLO_MOUSE
9	545	10.7	603	1	POLO_MOUSE
10	537	10.6	603	1	POLO_MOUSE
11	536	10.6	603	1	POLO_MOUSE
12	526.5	10.4	1518	1	POLO_MOUSE
13	514.5	10.1	683	1	POLO_MOUSE
14	508.5	10.0	403	1	POLO_MOUSE
15	504.5	9.9	632	1	POLO_MOUSE
16	503.5	9.9	397	1	POLO_MOUSE
17	501.5	9.9	395	1	POLO_MOUSE
18	498	9.8	615	1	POLO_MOUSE
19	498	9.8	661	1	POLO_MOUSE
20	496.5	9.8	649	1	POLO_MOUSE
21	496.5	9.8	649	1	POLO_MOUSE
22	496.5	9.8	649	1	POLO_MOUSE
23	493.5	9.7	1142	1	POLO_MOUSE
24	490.5	9.6	705	1	POLO_MOUSE
25	489.5	9.6	343	1	POLO_MOUSE
26	487.5	9.6	384	1	POLO_MOUSE
27	486	9.6	1037	1	POLO_MOUSE
28	485.5	9.6	344	1	POLO_MOUSE
29	479.5	9.4	345	1	POLO_MOUSE
30	477.5	9.4	774	1	POLO_MOUSE
31	471.5	9.3	282	1	POLO_MOUSE
32	469	9.2	367	1	POLO_MOUSE
33	468	9.2	1081	1	POLO_MOUSE

34	467.5	9.2	752	1	MRK4_HUMAN	Q96134 homo sapien
35	467	9.2	611	1	SNR1_CANGA	Q00372 candida gla
36	465.5	9.2	309	1	AURC_HUMAN	Q9ub99 homo sapien
37	462.5	9.1	633	1	SNR1_YEAST	P06782 saccharomyc
38	454.5	9.0	556	1	PDPK_HUMAN	O15530 homo sapien
39	454	8.9	559	1	PDPK_RAT	O55173 rattus norv
40	453.5	8.9	620	1	SNR1_CANAL	P52497 candida alb
41	450	8.9	559	1	PDPK_MOUSE	Q92240 mus musculu
42	449.5	8.9	794	1	KIT1_HUMAN	Q81d03 homo sapien
43	447	8.8	776	1	MRK3_HUMAN	P27448 homo sapien
44	445	8.8	502	1	KIT1_SECCR	Q02723 secale cere
45	441.5	8.7	533	1	KCCD_RAT	P15791 rattus norv

## ALIGNMENTS

RESULT 1	ID	POLO_DROME	STANDARD	PRT	576 AA.
AC	P52304	Q9VWB2			
DT	01-OCT-1996 (Rel. 34, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
DE	Protein kinase polo (EC 2.7.1.-).				
GN	POLO OR CG12306.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Canton-S;				
RX	MEDLINE=92084090; PubMed=1660828;				
RA	Llamaras S., Moreira A., Tavares A., Girdham C., Spruce B.A.,				
RA	Gonzalez C., Kares R.E., Glover D.M., Sunkel C.E.;				
RT	"Polo encodes a protein kinase homolog required for mitosis in				
RT	Drosophila.";				
RL	Gene Dev. 5:2153-2165 (1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Berkley;				
RX	MEDLINE=20196006; PubMed=10731132;				
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Mananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,				
RA	Stutson G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,				
RA	Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abdel J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,				
RA	Baillet R.M., Basu A., Baxendale J., Bayraktaroglu U., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,				
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,				
RA	Fosler A., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaeser K.,				
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,				
RA	Jalali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				
RA	Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,				
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,				
RA	Merklov G., Mlshina N.V., Mobaric C., Morris J., Moshrefi A.,				
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,				
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacht J.M.,				
RA	Palazzo M., Pittman G.S., Pan S., Pollard W., Puti V., Reese M.G.,				
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,				
RA	Shue B.C., Siden-Klamos I., Simpson W., Skupski M.P., Smith T.,				

CC	EMBL; X63361; CAA44963.1; -.
DR	EMBL; AE003514; AAF49036.1; -.
DR	PIR; S22127; S22127.
DR	HSSP; O63450; 1A06.
DR	FLYBase; FBgn0003124; polo.
GO	GO:0005813; C:centrosome, IDA.
GO	GO:0005819; C:spindle, IDA.
GO	GO:0004674; F:protein serine/threonine kinase activity, IDA.
GO	GO:0007140; P:male meiosis, IMP.
GO	GO:0007067; P:mitosis, IMP.
InterPro	IPR000959; POLO box.
InterPro	IPR000719; Prot_kinase.
InterPro	IPR008271; Ser_thr_kin_AS.
InterPro	IPR002290; Ser_thr_pkinase.
DR	Pfam; PF00659; Pkinase; 1.
DR	Pfam; PF00659; POLO_box; 2.
DR	ProDom; PD000001; Prot_kinase; 1.
DR	SMART; SMO0220; S_TKC; 1.
DR	PROSITE; PS002078; POLO_BOX; 2.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR	PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW	Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat.
FT	DOMAIN 25 277
FT	ND BIND 31 39
FT	BP BIND 54 54
FT	ACT SITE 148 148
FT	DOMAIN 398 461
FT	DOMAIN 496 564
FT	CONFLICT 187 187
SEQUENCE	576 AA; 66973 MW; 5022B9AC0E888FAD CRC64; P -> A (IN REF. 1).

QY	176	NYISPEATSAHGLESVDWISLQCFEYTLILGRPEFDITVKNTLNRVLADYEMPSFLS	2359
Db	188	NYAPLELTITKKGHSFEVDIWSIGCWVYTLVLVQGPPEFTKTLMDITYSKI KKECEYVPSYLR	247
QY	236	IEAKDILHOLRNPDRILSSSVLDHPMS-----RNSSTKSDLG--TYVDSI	2833
Db	248	KPAADVIATMLQNPESRPAIQOLNFEELKSKVPMELPSSCLITMAPRIGSDNTIYDSM	3070
QY	284	D-----SG-----HATISATYASGTSISGSLPDKRLLIGQLPDKMT	3633
Db	308	HRKPLMENGIRPDTRLESTFLKXNLMDAITASQVC-----	3465
QY	324	VPRKAKSSTDPSSSGSGNSFTYQMGNOETSNSGRGVIG	362
Db	346	-----RHSESDYRS--DIESLYQOLTN--LINGPRILQ	374

```
CC CC      RESULT_2  
CC ID       CNK_MOUSE     STANDARD;          PRT;           631 AA.  
AC Q60806; Q60822; Q9R009;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Cytochrome-inducible serine/threonine-protein kinase (EC 2.7.1.37) [PGF  
DE Inducible Kinase].  
GN CNK OR FNK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN=NIH Swiss;  
RX MEDLINE=95247749; PubMed=7730342;  
RA Donohue P.U., Alberts G.F., Guo Y., Winkles J.A.;  
RT "Identification by targeted differential display of an immediate early  
RL gene encoding a putative serine/threonine kinase." ;  
RU J. Biol. Chem. 270:10351-10357(1995).  
LN [2]  
RP SEQUENCE OF 333-437 FROM N.A. (ISOFORM 2).  
RC STRAIN=NH Swiss;  
RA Kauselmann G., Weiler M., Kuhl D. ;  
RL Submitted (NOV-1999) to the EMBL/genbank/DDBJ databases.  
RN [3]  
RP FUNCTION, AND PHOSPHORYLATION.  
RX CHEDLINE=98343954; Pubmed=9677325;  
RA Maese D., Feng Y., Hanselew B., Winkless J.A., Longo D.L., Ferris D.K.;  
RT "Expression and phosphorylation of fibroblast-growth-factor-inducible  
kinase (Fnk) during cell-cycle progression.";  
RL Biochem. J. 333:655-660(1998).  
CC -!- FUNCTION: Serine/threonine protein kinase involved in regulating M  
phase functions during the cell cycle. May also be part of the  
signaling network controlling cellular adhesion. In vitro, is able  
to phosphorylate CDC25C and casein (By similarity).  
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -!- SUBUNIT: Binds to the calcium/integrin-binding protein (CIB). This  
interaction probably occurs via the POLO-box domain (By  
similarity).  
CC -!- SUCCELLULAR LOCATION: Membrane-associated (By similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC   Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC   IsoId=Q60806-1; Sequence=Dlsplayed;  
CC Name=2;  
CC   IsoId=Q60806-2; Sequence=VSP_004927;  
CC -!- TISSUE SPECIFICITY: Expressed in skin.  
CC -!- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS  
cells exit mitosis.  
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
CC -!- SIMILARITY: Contains 2 POLO box domains.
```





Best Local Similarity 33.9%; Pred. No. 3,1e-24;  
Matches 126; Conservative 64; Mismatches 121; Indels 61; Gaps 7;

QY 15 GNLLKGGFAGVYRAESHTGLVLAIKMIDKKAMKAMQVONENKICOLKHPSTLE 74  
DB 34 GRLLKGGFAGVYRAESHTGLVLAIKMIDKKAMKAMQVONENKICOLKHPSTLE 74  
QY 75 LYNPEDESNVYLVLEMCHGEMNRVYLNKRVKPSSENAHFMHIOITGMILYHSHGILH 134  
DB 94 FSHHEADNDIYILELCSRSKSLAHMKAR-HTLLEPVRKYKOLISGLKYLHORGILH 152  
QY 135 RDLTSLNLLTRNNKIKIADFGATQOLKMPHEKHYLLCGPNYISPEIATRSAGHESDV 194  
DB 153 RDLKLGFFITDNNMELKGVGLAARLEPQRKKTICGPNVYAPVLAHQHGPPADV 212  
QY 195 WSLGCMFFTLIGRPPTDTPVNTLKNVYLAADVEMSPFISIAKDLIHLRRNPADRL 254  
DB 213 WSLGCMFFTLIGRPPTDTPVNTLKNVYLAADVEMSPFISIAKDLIHLRRNPADRL 254  
QY 255 SLSSVLDHPMRSNSTKSKDQGVEDSIDSGHATITATPASSST-----SIGG 304  
DB 273 SIEQILHDFEFTK-----GYTPRL-----PVSSCTVVDLPNPAASLFAKATK 318  
QY 305 SLTPDKR-----RLLIGQP-----LPNKMVFPNNKST----- 332  
DB 319 SLFGRRKSKNNHSEEDNVSLVSGIMRTSIGHDPVPPAPASALAPVSLVETAEDS 378  
QY 333 -----DFSSSGDG 340  
DB 379 SPRGTASSGDG 390

RESULT 4  
CNK\_HUMAN  
ID Q9H4B4; Q15767; STANDARD; PRT; 646 AA.  
AC 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Cytokine-inducible serine/threonine-protein kinase (EC 2.7.1.37) (FCF-inducible kinase) (Proliferation-related kinase).  
GN CNK OR FNK OR PRK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=20493044; PubMed=11039900;  
RA Kauselmann G., Rehl M., Andreesen R., Kautmann M., Kuhl D., Strubhardt K.;  
RT "Addression induced expression of the serine/threonine kinase Fnk in human macrophages.";  
RL Oncogene 19:4832-4839 (2000).  
RN [2]  
RP SEQUENCE OF 28-646 FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=96325053; PubMed=8702627;  
RA Li B., Ouyang B., Pan H., Reissmann P.T., Slamon D.J., Arceci R., Lu L., Dai W.;  
RT "Ptk, a cytokine-inducible human protein serine/threonine kinase whose expression appears to be down-regulated in lung carcinomas.";  
RL J. Biol. Chem. 271:19402-19408 (1996).  
RN [3]  
RP CHARACTERIZATION.  
RX MEDLINE=96019242; PubMed=9353331;  
RA Ouyang B., Pan H., Lu L., Li J., Stambrook P., Li B., Dai W.;  
RT "Human Ptk is a conserved protein serine/threonine kinase involved in regulating M phase functions.";  
RL J. Biol. Chem. 272:28646-28651 (1997).  
RN [4]

RP CHARACTERIZATION.  
RX MEDLINE=20027391; PubMed=10557092;  
RA Ouyang B., Li W., Pan H., Meadows J., Hoffmann I., Dai W.;  
RT "The physical association and phosphorylation of Cdc25C protein phosphatase by Ptk.";  
RL Oncogene 18:6029-6036 (1999).  
CC -1- FUNCTION: Serine/threonine protein kinase involved in regulating M phase functions during the cell cycle. May also be part of the signaling network controlling cellular adhesion. In vitro, is able to phosphorylate Cdc25C and casein.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -1- SUBUNIT: Binds to the calcium/integrin-binding protein (CIB). This interaction probably occurs via the POLO-box domain.  
CC -1- SUBCELLULAR LOCATION: Membrane-associated.  
CC -1- TISSUE SPECIFICITY: Transcripts are highly detected in placenta, lung, followed by skeletal muscle, heart, pancreas, ovaries and kidney and weakly detected in liver and brain. May have a short half-life. In cells of hematopoietic origin, strongly and exclusively detected in terminally differentiated macrophages. Transcript expression appears to be down-regulated in primary lung tumor.  
CC -1- INDUCTION: Cytokine and cellular adhesion trigger FNK induction.  
CC -1- PTM: PHOSPHORYLATED AS CELLS ENTER MITOSIS AND DEPHOSPHORYLATED AS CELLS EXIT MITOSIS (BY SIMILARITY).  
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. CDC5/Polo subfamily.  
CC -1- SIMILARITY: Contains 2 POLO box domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.1sb-sib.ch/announce/> or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
CC -----  
DR EMBL; AJ293866; CAC10659.1; -.  
DR EMBL; U56998; AAC50637.1; ALT\_INT.  
DR Genbank; HGNC:2154; CNK.  
DR GK; Q9H4B4; -.  
DR MIM; 602913; -.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.  
DR GO; GO:0000074; P:regulation of cell cycle; TAS.  
DR InterPro; IPR000959; POLO\_box.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00659; POLO\_box; 2.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR PROSITE; PSS0078; POLO\_BOX; 2.  
DR PROSITE; PSS0107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PSS0108; PROTEIN\_KINASE\_ST; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat; Phosphorylation.  
FT DOMAIN 62 314 PROTEIN KINASE.  
FT NP\_BIND 68 76 ATP (BY SIMILARITY).  
FT BINDING 91 91 ATP (BY SIMILARITY).  
FT ACT\_SITE 185 185 BY SIMILARITY.  
FT DOMAIN 470 537 POLO\_BOX 1.  
FT DOMAIN 567 637 POLO\_BOX 2.  
FT CONFLICT 99 99 V -> A (IN REF. 2).  
FT CONFLICT 353 353 V -> D (IN REF. 2).  
FT CONFLICT 419 419 H -> D (IN REF. 2).  
FT CONFLICT 464 470 FSEWGF -> VSKWVDY (IN REF. 2).  
FT CONFLICT 522 522 R -> P (IN REF. 2).  
SQ SEQUENCE 646 AA; 71789 MW; C20147CD0FCA3B4 CRC64;

Query Match 11.6%; Score 589.5; DB 1; Length 646;  
Best Local Similarity 37.5%; Pred. No. 3.7e-24;

Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;

QY 15 GNLGKGSFAGYVRAESITHTGLEVAIKMIDKAMRYAGVQVQVNEVKIHQCKHPSILE 74  
 DB 65 GRLGKGGRFARCYEADTETGSAVAVKVIPOSRVVPHQREKLIETELHRLQHHIVR 124  
 QY 75 LYNYPEDSNVYVLYVEMCHNGEMNRYLKRVKVPFSENEARHPHQIITGMLYLSHGILH 134  
 DB 125 FSHHFEADANITYIFELCSKRSKLAHWKAR-HTLEPEVRYIRQLVSGKILHOGGILH 183  
 QY 135 RDLTSLNLLTTRNNMIKIDFGIATQLKMPHEKHYTLCTPNYISPEIATRSAGLESVD 194  
 DB 184 RDLKGNFPTINEMELKVDGFLAARLEPDEQKKTICGTPNVVADEVLLRQHGGEADV 243  
 QY 195 NSLGCFFYTLIGRPPFDYTKNTLANKVYLADYEMPSFLSIAKOLHQLRRNADRL 254  
 DB 244 NSLGCVMYTLIGRPPFDYTKNTLANKVYLADYEMPSFLSIAKOLHQLRRNADRL 303  
 QY 255 SLSSVLDHPFMRNSTKSDGTVEDSIDSGHATISTATISSST-----SISG 304  
 DB 304 SLDQILRHDFPTK-----GYTPDRL-----PISCVTPDLPNPANSLPAKVTX 349  
 QY 305 SLFDKRR 311  
 DB 350 SLFVRKK 356

RESULT 5  
 SNK\_MOUSE STANDARD; PRT; 682 AA.

ID SNK\_MOUSE  
 AC P1351;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Serine/threonine-protein kinase SNK (BC 2.7.1.-) (Serum inducible kinase).  
 GN SNK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92375085; PubMed=1508211;  
 RA Simmons D.L., Neel B.G., Stevens R., Ewelt G., Ertkeon R.L.;  
 RT "Identification of an early-growth-response gene encoding a novel putative protein kinase."  
 RL Mol. Cell. Biol. 12:4164-4169 (1992).  
 CC -1- FUNCTION: May play a role in the division of at least some cell types, such as fibroblasts, and could function in embryogenesis, wound healing or neoplasia.  
 CC -1- TISSUE SPECIFICITY: Brain, lung and heart.  
 CC -1- INDUCTION: By serum and phorbol ester.  
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
 CC CDCs/Polo subfamily.  
 CC -1- SIMILARITY: Contains 2 POLO box domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.1sb-sib.ch/announce/ or send an email to 1license@sib-sib.ch).  
 CC EMBL; M96163; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A44493; A44493.  
 DR MGD; MGI:1099780; Snk.  
 DR InterPro; IPR000959; POLO\_box.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; I.

DR Pfam; PF00659; POLO\_box; 2.  
 DR Prodom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS50078; POLO\_BOX; 2.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat.  
 FT DOMAIN 54 59 POLY-HIS  
 FT BINDING 79 331 PROTEIN\_KINASE.  
 FT ACT\_SITE 108 108 ATP (BY SIMILARITY).  
 FT ACT\_SITE 202 202 BY SIMILARITY.  
 FT DOMAIN 507 570 POLO\_BOX\_1.  
 FT DOMAIN 603 674 POLO\_BOX\_2.  
 SQ SEQUENCE 682 AA; 77811 MW; 586DEAFD7208A9D CRC64;

Query Match 11.1%; Score 561.5; DB 1; Length 682;  
 Best local Similarity 41.6%; Pred. No. 1.2e-22;  
 Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

QY 15 GNLGKGSFAGYVRAESITHTGLEVAIKMIDKAMRYAGVQVQVNEVKIHQCKHPSILE 74  
 DB 82 GYVLGKGGRFARCYEADTETGSAVAVKVIPOSRVVPHQREKLIETELHRLQHHIVR 141  
 QY 75 LYNYPEDSNVYVLYVEMCHNGEMNRYLKRVKVPFSENEARHPHQIITGMLYLSHGILH 134  
 DB 142 FYHFEADANITYIFELCSKRSKLAHWKAR-KVLEPEVRYIRQLVSGKILHOGGILH 200  
 QY 135 RDLTSLNLLTTRNNMIKIDFGIATQLKMPHEKHYTLCTPNYISPEIATRSAGLESVD 194  
 DB 201 RDLKGNFPTINEMELKVDGFLAARLEPDEQKKTICGTPNVVADEVLLRQHGGEADV 260  
 QY 195 NSLGCFFYTLIGRPPFDYTKNTLANKVYLADYEMPSFLSIAKOLHQLRRNADRL 254  
 DB 261 NSLGCVMYTLIGRPPFDYTKNTLANKVYLADYEMPSFLSIAKOLHQLRRNADRL 320  
 QY 255 SLSSVLDHPF 264  
 DB 321 SLDQILRHDF 330

RESULT 6  
 SNK\_MOUSE STANDARD; PRT; 682 AA.

ID SNK\_MOUSE  
 AC Q9R012;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Serine/threonine-protein kinase SNK (BC 2.7.1.-) (Serum inducible kinase).  
 GN SNK.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxId=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9452760; PubMed=10523297;  
 RA Kaussidmann G., Weller M., Wulff P., Jessberger S., Konietzko U., Scafield J., Straub U., Berleiter-Hahn J., Streibhardt K., Kuhl D.;  
 RT "The polo-like protein kinases Pnk and Snk associate with a Ca(2+)-and integrin-binding protein and are regulated dynamically with synaptic plasticity."  
 RL EMBO J. 18:5528-5539 (1999).  
 CC -1- FUNCTION: May play a role in the division of at least some cell types, such as fibroblasts, and could function in embryogenesis, wound healing or neoplasia (By similarity).  
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
 CC CDCs/Polo subfamily.  
 CC -1- SIMILARITY: Contains 2 POLO box domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF136583; AAF08366.1; -  
 DR InterPro: IPR000959; POLO box.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00659; POLO box; 2.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR PROSITE: PS50078; POLO BOX; 2.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat.  
 FT DOMAIN 54 59 POLY-HIS.  
 FT BINDING 79 331 PROTEIN\_KINASE.  
 FT BINDING 85 93 ATP (BY SIMILARITY).  
 FT ACT\_SITE 108 108 ATP (BY SIMILARITY).  
 FT DOMAIN 202 202 BY SIMILARITY.  
 FT DOMAIN 570 570 POLO BOX 1.  
 FT DOMAIN 603 674 POLO BOX 2.  
 SQ SEQUENCE 682 AA; 77919 MW; 58C50DBDBE83D5F3 CRC64;

Query Match 11.0%; Score 560.5; DB 1; Length 682;  
 Best Local Similarity 41.6%; Pred. No. 1.4e-22;  
 Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

QY 15 GNLLGKSGFAGYRRASIHITGLEVATKMDKKAMKAGVQVQNVKTHCQKHSILE 74  
 DB 82 GVLGKSGFAGYRRASIHITGLEVATKMDKKAMKAGVQVQNVKTHCQKHSILE 74  
 QY 75 LYNVEDSNVYVLEMCNCHGEMNRVYKRVKPFSENEARHFMHQITGLYLHSHGILH 134  
 DB 142 FHYHEDKENIYILLECYGRSMAMHILKAR-KVLEPEVRYVLRQVSGKYLHBEELIH 200  
 QY 135 RLUTLSNLLTTRNNKIKADFGLATOLKMPHEKHVTLCTPVYISPELATRSAGLESV 194  
 DB 201 RLKIKGNFPIINAMEIKVDFGLAARLEBIEHRRRTICGTPYVLSPEVANKGHCESPI 260  
 QY 195 WSLGCMFVYLLIGRPEDFDVYKNTLKVLAADYEMPSFLSTEARDLIHQLRRNPADRL 254  
 DB 261 WALGCMVMTLIGRPFETTNLKYTRCIRBARVTPSSLAPAKHLIASMLSKNEDRP 320  
 QY 255 SLSSVLDHPF 264  
 DB 321 SLDDIIRHDF 330

RESULT 7  
 ID SNK\_HUMAN STANDARD; PRT; 685 AA.  
 AC Q9NXY3; O60679; Q96CV7; Q9UE61;  
 DT 16-OCT-2001 (Rel. 40. Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serine/threonine-protein kinase SNK (BC 2.7.1.-) (Serum inducible  
 DE kinase).  
 GN SNK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN NCBI\_Taxid=9606;  
 RP SEQUENCE FROM N.A.  
 RA Ouyang B., Dai W.;  
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon;  
 RA Anderson K.M., Nerurkar S.S., Hansbury M.J., Formwald J., Scott G.,  
 RA Bouzyk M., Mui P., Imbruglia C.S., Carlson K., Marshall L.A.,  
 RA Koshak A.K.;  
 RT "Identification and characterization of human serum-inducible kinase  
 RT (SNK), a novel member of the polo-kinase family of cell cycle  
 RT regulators: potential implication for regulation of vascular smooth  
 RT muscle proliferation.";  
 RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skin;  
 RX MEDLINE=2238257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheele C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stalderon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Tshibuyi S., Cantoni P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollath S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Buterfield Y.S., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [4]  
 RP SEQUENCE OF 110-408 FROM N.A.  
 RA Fidler C., Boulwood J., Wang Jabs E., Waincoat J.S.;  
 RL Submitted (Jan-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: May play a role in the division of at least some cell  
 CC types, such as fibroblasts, and could function in embryogenesis,  
 CC wound healing or neoplasia (By similarity).  
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
 CC -1- SIMILARITY: Contains 2 POLO box domains.  
 CC This SNK-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF059617; AAC14573.1; -  
 DR EMBL: AF223574; AAF62897.1; -  
 DR EMBL: U85755; AAD00575.1; -  
 DR EMBL: BC013879; AAH13879.1; -  
 DR MIM: 607023; -  
 DR InterPro: IPR000959; POLO box.  
 DR InterPro: IPR000719; Prot\_Kinase.  
 DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00659; POLO box; 2.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR PROSITE: PS50078; POLO BOX; 2.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat.  
 FT DOMAIN 57 64 POLY-HIS.

```

CC FT DOMAIN 82 334 PROTEIN KINASE.
CC FT NP_BIND 88 96 ATP (BY SIMILARITY).
CC FT BINDING 111 111 ATP (BY SIMILARITY).
CC FT ACT_SITE 205 205 BY SIMILARITY.
CC FT DOMAIN 510 573 POLO BOX 1.
CC FT DOMAIN 606 677 POLO BOX 2.
CC FT CONFLICT 28 28 A -> G (IN REF. 1).
CC SEQ SEQUENCE 685 AA; 78236 MW; 6429P6FERD30B33 CRC64;

Query Match 11.0%; Score 560.5; DB 1; Length 685;
Best Local Similarity 41.6%; Pred. No. 1,4e-22;
Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

QY 15 GNLGKGFAGYVRAESIHITGLEVAIKMDKKAMKAGVQVRQNVKIKHCOLKHSIIE 74
DB 85 GVLGKGFAGYVRAESIHITGLEVAIKMDKKAMKAGVQVRQNVKIKHCOLKHSIIE 144
QY 75 LYNVEDSNVYVLEMCNCHGENRKYKRVKPFSENEARHFMHQITITGMLYHSHGIIH 134
DB 145 FHYHEDKENIYILLEYCSRSAHLKAR-KVLTEPRVYVRLQVSGIKYHGEIILH 203
QY 135 RDLVLSNLLITNNMIKIDFGIATOLKMPHEKHYLCGTPNTISPEIATRSAGLESVD 194
DB 204 RDLKLGKGFAGYVRAESIHITGLEVAIKMDKKAMKAGVQVRQNVKIKHCOLKHSIIE 263
QY 195 NSLGGCFYVLLIGRPFEDDTVKNLKVADYEMPSFLISAKDLIHQLLRNADRL 254
DB 264 WALGCMYVLLIGRPFEDDTVKNLKVADYEMPSFLISAKDLIHQLLRNADRL 323
QY 255 SLSSVLDHPF 264
DB 324 SLDDIRHDF 333

RESULT 8
STK6_XENLA STANDARD; PRT; 407 AA.
ID STK6_XENLA
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine protein kinase Eg2 (EC 2.7.1.37) (p46Eg265).
GN Eg2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCB1_TaxID=8355;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Egg;
RX MEDLINE=98122855; PubMed=9454730;
RA Roghi C., Giet R., Uzbekov R., Morin N., Chatterain I., Le Guellec R.,
RA Couturier A., Doree M., Philippe M., Prigent C.;
RT "The Xenopus protein kinase p62 associates with the centrosome in a
RT cell cycle-dependent manner, binds to the spindle microtubules and is
RT involved in bipolar mitotic spindle assembly."
RT J. Cell Sci. 111:557-572(1998).
CC -1- FUNCTION: Associates with the centrosome in a cell-cycle dependent
CC manner and invades the microtubules at the poles of the spindle
CC during mitosis suggesting that it may be involved in the correct
CC formation of bipolar mitotic spindles.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBCELLULAR LOCATION: Localizes to the spindle pole during mitosis
CC especially from prophase through anaphase. Partially colocalised
CC with gamma tubulin in the centrosome, from S to M phase.
CC -1- TISSUE SPECIFICITY: Highly expressed in ovary and testis.
CC -1- PTM: Phosphorylated.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC Aurora subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 217207; CAA78915.1; -.
CC PIR; S52243; S52243.
CC DR HSSP; P24941; 1A01.
CC DR InterPro; IPR000719; Prot_kinase.
CC DR InterPro; IPR008271; Ser_thr_kin_AS.
CC DR InterPro; IPR002290; Ser_thr_kinase.
CC DR Pfam; PF00069; pkinase; 1.
CC DR ProDom; PD000001; Prot_kinase; 1.
CC DR SMART; SM00220; S_TKc; 1.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC DR Cell cycle; Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC Phosphorylation.
CC FT DOMAIN 140 390 PROTEIN KINASE.
CC FT NP_BIND 146 154 ATP (BY SIMILARITY).
CC FT BINDING 169 169 ATP (BY SIMILARITY).
CC FT ACT_SITE 263 263 BY SIMILARITY.
CC SEQ SEQUENCE 407 AA; 46372 MW; DE1628A2CD11277 CRC64;

Query Match 10.9%; Score 554.5; DB 1; Length 407;
Best Local Similarity 39.7%; Pred. No. 1,4e-22;
Matches 106; Conservative 60; Mismatches 98; Indels 3; Gaps 2;

QY 9 IEDFKVGNLIGKGFAGYVRAESIHITGLEVAIKMDKKAMKAGVQVRQNVKIKHCOLK 68
DB 137 LEDFEIGRLGKGFAGYVRAESIHITGLEVAIKMDKKAMKAGVQVRQNVKIKHCOLK 196
QY 69 HPSILELYNVEDSNVYVLEMCNCHGENRKYKRVKPFSENEARHFMHQITITGMLYH 128
DB 197 HPNILLVGYFHDASHRVYLLDYAPGSELFRELOKCTR-FDDQRSAMYIKQLAEALLYCH 255
QY 129 SHGIIHRDLTSLNLLITNNMIKIDFGIATOLKMPHEKHYLCGTPNTISPEIATRSAG 188
DB 256 SKKVHNRDIPKPELILGSGNGLKIDFGWS--VHABSSRTLLCGTLVLPFPMIGRMH 313
QY 189 GLESDVNSLSCMFYVLLIGRPFEDDTVKNLKVADYEMPSFLISAKDLIHQLLRN 248
DB 314 DEIVDMSLGVLCYERLVGKPFEDDTVKNLKVADYEMPSFLISAKDLIHQLLRN 373
QY 249 NPADRSLSSVLDHPFMSNNSSTKSD 275
DB 374 NPVHRLPLKGVLEHPWIKNSQLKKD 400

RESULT 9
STK6_XENLA STANDARD; PRT; 408 AA.
ID STK6_XENLA
AC 091819;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine protein kinase Eg2-like (EC 2.7.1.37) (p46XLEg22).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCB1_TaxID=8355;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Egg;
RX MEDLINE=98122855; PubMed=9454730;
RA Roghi C., Le Guellec R., Paris J., Couturier A., Philippe M.;
RT "Eg2, selected by differential screening encodes a new Xenopus protein
RT kinase family."
RT Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBCELLULAR LOCATION: Localizes to the spindle pole during mitosis

```

especially from prophase through anaphase. Partially colocalised with gamma tubulin in the centrosome, from S to M phase (by similarity).

-1- PFM: Phosphorylated (By similarity).

-1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. Aurora subfamily.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: Z17206; CAA78914.1; ALT\_INIT.

DR HSSP: P24941; 1A01.

DR InterPro: IPR000719; Prot kinase.

DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.

DR InterPro: IPR002290; Ser\_thr\_pkinase.

DR InterPro: IPR001245; Tyr\_pkinase.

DR Pfam: PF00069; pkinase; 1.

DR ProDom: PD000001; Prot\_kinase; 1.

DR SMART: SM00220; S\_TKc; 1.

DR SMART: SM00219; TyKc; 1.

DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.

DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.

DR PROSITE: PS00108; PROTEIN KINASE\_ST; 1.

CC Cell cycle; Transferrase; Serine/threonine-protein kinase; ATP-binding; Phosphorylation.

KW Phosphorylation.

FT DOMAIN 140 390 PROTEIN KINASE.

FT NP\_BIND 146 154 ATP (BY SIMILARITY).

FT BINDING 169 169 ATP (BY SIMILARITY).

FT ACT\_SITE 263 263 BY SIMILARITY.

SQ SEQUENCE 408 AA; 46461 MW; 97F6A60C7357A8E8 CRC64;

Query Match 10.8%; Score 548; DB 1; Length 408;

Best Local Similarity 39.9%; Pred. No. 3; 2e-22;

Matches 107; Conservative 60; Mismatches 97; Indels 4; Gaps 3;

QY 9 IEDPKVNLIGKSGFAGVYRAESIHGTGLVATKMDKKMYKAGVQRYQNVKIHQOLK 68

DB 137 LDDFELGRGLKPKNGVTLARERBEKFTALKVLFKSLDEKAGVHQRREVEIQSHLR 196

QY 69 HPSILELVYFEDSNVYVLEMCNENRRLKRVKPFSENEARHFMHQITGMLYLH 128

DB 197 HNNIILYGFYFDASRVYILDYAPGSELFRELOKCTR-FDQSRALYIKQIAEMLLYCH 255

QY 129 SHGIIHRDLTSLNLTTRMNKIKADFGIATQIKMPHEKHYLCSTPNYISPEIATRSAN 188

DB 256 SKKVYHRDIPKPNLLSGNGELKIADFGWS--VHAPSSRRITLCSTLDYLPPEMTEGRMH 313

QY 189 GIESDWSISGCMFYTLIGRPEDTDYVNTLKVLAIDYEMSPFLEAKLIHQILRR 248

DB 314 DEKVLMSGLVCYERLIVGKPPFDTHQETVRRISKVEFGYPPYVSEAKDLVSKLKH 373

QY 249 NPADRLSSVYLDHPMRSNS-STYSKD 275

DB 374 NPNHRLPLKGVLEHPIVANSQOPKKD 401

RESULT 10

PK1\_MOUSE STANDARD; PRT; 603 AA.

AC Q07832;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Serine/threonine-protein kinase PLK (BC 2.7.1.-) (PLK-1) (Serine-threonine protein kinase 13) (STPK13).

GN PLK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxId=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Bone marrow;

RX MEDLINE=93281660; PubMed=8099445;

RA Clay F.J., McEwen S.J., Bertonecello I., Wilks A.F., Dunn A.R.;

RT "Identification and cloning of a protein kinase-encoding mouse gene, Plk, related to the polo gene of Drosophila."

RL Proc. Natl. Acad. Sci. U.S.A. 90:4882-4886(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6 X CBA; TISSUE=Thymus;

RX MEDLINE=94289293; PubMed=8018557;

RA Hamanaka R., Malold S., Smith M.R., O'Connell C.D., Longo D.L., Ferris D.K.;

RT "Cloning and characterization of human and murine homologues of the Drosophila polo serine-threonine kinase."

RL Cell Growth Differ. 5:249-257(1994).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=94067140; PubMed=7902533;

RA Lake R.J., Jelinek W.R.;

RT "Cell cycle- and terminal differentiation-associated regulation of the mouse mRNA encoding a conserved mitotic protein kinase."

RL Mol. Cell. Biol. 13:7793-7801(1993).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB/N; TISSUE=Mammary gland;

RX MEDLINE=22388257; PubMed=12477932;

RA Strubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins P.S., Wagner L., Sherman C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullany S.J., Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gamarate P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -1- FUNCTION: May be required for cell division and may have a role during G1 or S phase.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- TISSUE SPECIFICITY: Newborn and adult spleen, fetal and newborn kidney, liver, brain, thymus and adult bone marrow, thymus, ovary and testes.

CC -1- DEVELOPMENTAL STAGE: In the thymus, levels increased during fetal development, were highest in newborn animals and decreased in the adult. In the testes, the PLK levels were higher in the adult than in prepubescent mice, while in the ovary, the levels were higher in the prepubescent mice. Accumulates to a maximum during the G2 and M phases, declines to a nearly undetectable level following mitosis and throughout G1 phase, and then begins to accumulate again during S phase.

CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.

CC -1- CC05/Polo subfamily.

CC -1- SIMILARITY: Contains 2 POLO box domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L06144; AAA59948.1; -  
 DR EMBL; U01063; AAA56635.1; -  
 DR EMBL; L19558; AAA16071.1; -  
 DR EMBL; BC006880; AAA06880.1; -  
 DR PIR; A47545; A47545.  
 DR PIR; A54596; A54596.  
 DR MGD; MG1; 97621; PK.  
 DR InterPro; IPR000959; POLO\_box.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR Pfam; PF00659; PKinase; 1.  
 DR Pfam; PF00659; POLO\_box; 2.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PSS0078; POLO\_BOX; 2.  
 DR PROSITE; PSS0107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PSS0108; PROTEIN\_KINASE\_ST; 1.  
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;  
 KM Nuclear protein.  
 FT DOMAIN 53 305 PROTEIN KINASE.  
 FT NP\_BIND 59 67 ATP (BY SIMILARITY).  
 FT BINDING 82 82 ATP (BY SIMILARITY).  
 FT ACT\_SITE 176 176 BY SIMILARITY.  
 FT DOMAIN 417 480 POLO BOX 1.  
 FT DOMAIN 515 584 POLO BOX 2.  
 FT CONFLICT 4 4 A -> V (IN REF. 1).  
 FT CONFLICT 15 15 A -> T (IN REF. 1).  
 FT CONFLICT 23 23 P -> L (IN REF. 1).  
 FT CONFLICT 27 27 P -> A (IN REF. 1).  
 FT CONFLICT 29 29 G -> S (IN REF. 1).  
 FT CONFLICT 41 41 P -> L (IN REF. 1).  
 FT CONFLICT 54 54 V -> I (IN REF. 1).  
 FT CONFLICT 495 495 A -> R (IN REF. 1).  
 SQ SEQUENCE 603 AA; 68300 MW; 1B980646366EFA10 CRC64;

Query Match 10.7%; Score 545; DB 1; Length 603;  
 Best Local Similarity 27.2%; Pred. No. 7, 8e-22;  
 Matches 152; Conservative 95; Mismatches 190; Indels 122; Gaps 20;

QY 15 GNLLGKSGFAGVYRAESIHTEVAIKMIDKKAMVAGVQVONEVKIHQOLKPSILE 74  
 DB 56 GRFLGKSGFAGVYRAESIHTEVAIKMIDKKAMVAGVQVONEVKIHQOLKPSILE 115  
 QY 75 LKMYFSDSVYVYIVEMCHNGENKRYLKRVPFSENEARHPHQIITGMLYHSHGILH 134  
 DB 116 FHDFFSDSPFVFAVLELCRRRSLLLELHKR-KALTEPEAAVYRQIVLGGQYHNRQVH 174  
 QY 135 RDLTSLNLLTRNMNLIKADFGATOLKMPHEKHYLTGPNYISPEIATRSNAGLESYV 194  
 DB 175 RDLKLGMLNLEBLEVKIGDFGATVVEGEKSKTLCTGPNYIAEVLVSKSHSEVDV 234  
 QY 195 WSLGCMFYTLTIGRPFDVDTVNTLVKVLADYEMPSFLISPAKOLIHOLLRRNPADRL 254  
 DB 235 WSLGCMFYTLTIGRPFDVDTVNTLVKVLADYEMPSFLISPAKOLIHOLLRRNPADRL 294  
 QY 255 SLSSVLDHFMNSNSTKSKDCTVEDSIDSGHATIS---TATASSNSISGLPD-K 309  
 DB 295 TLHLELNDLFT-----SGYIPARLPITCLTIPPRSIAPSSLDPSG 336  
 QY 310 RRLT-----IGQPLNKMVTFPPKNKSTDFSSSGDG-----NSFYTOMGNETSN-SGR 357  
 DB 337 RKLPLKVLNKGVEHPLDR-----PREKEPEVRETNAIECHSLDLQQLTVNASKPSER 392  
 QY 358 GRVYQDAEERPH-----SRVLRAYSSDRSGTS-----NSOSAKTYTMMER 398  
 DB 393 GLVRQEAEDPACIPFVWSKWD---YSDKYGLGYQLCDNSVGVLFNDSTRILTYL---- 445

QY 399 CHSAEMLSVSKSGGGENEERYSPDNNAN-----INPFKEKTS---SGSFERPD 448  
 DB 446 -NDGDSLOYIERDG-----TESVLTWSHNSLMKKTITLNYRNTWSEHLKAGANITR 500  
 QY 449 NNQALSNHLCPGKTPPPADPTPQETTVQWF-----GNLQIN----- 486  
 DB 501 EGDLELA-----RLPY-----LRTWFRTSAILHLNNGVQINFPDHTKIL 543  
 QY 487 AHLKRTTEYDSISPNRPFQ 505  
 DB 544 CPLMAAVTY--INERKDFQ 550

RESULT 11  
 PLK1\_RAT STANDARD; PRT; 603 AA.  
 ID PLK1\_RAT  
 AC 062673;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Serine/threonine-protein kinase PLK (EC 2.7.1.-) (PLK-1).  
 GN PLK.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RA Amstrup J., Hansen J.A., Hxirils Nielsen J.;  
 Submitted (MAY-1994) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: May be required for cell division and may have a role  
 CC during G1 or S phase (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
 CC CDC5/Polo subfamily.  
 CC -1- SIMILARITY: Contains 2 POLO box domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U0188; AAA1885.1; -  
 DR InterPro; IPR000959; POLO\_box.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR Pfam; PF00659; PKinase; 1.  
 DR Pfam; PF00659; POLO\_box; 2.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PSS0078; POLO\_BOX; 2.  
 DR PROSITE; PSS0107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PSS0108; PROTEIN\_KINASE\_ST; 1.  
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding; Repeat;  
 KM Nuclear protein.  
 FT DOMAIN 53 305 PROTEIN KINASE.  
 FT NP\_BIND 59 67 ATP (BY SIMILARITY).  
 FT BINDING 82 82 ATP (BY SIMILARITY).  
 FT ACT\_SITE 176 176 BY SIMILARITY.  
 FT DOMAIN 417 480 POLO BOX 1.  
 FT DOMAIN 515 584 POLO BOX 2.  
 SQ SEQUENCE 603 AA; 68313 MW; 107AFB3B7EDC002 CRC64;

Query Match 10.6%; Score 537; DB 1; Length 603;  
 Best Local Similarity 27.0%; Pred. No. 2, 1e-21;  
 Matches 151; Conservative 95; Mismatches 191; Indels 122; Gaps 20;

```

QY 15 GNLGKGSFAGVYRAESIHTGLEVALIKMIDKKAMKAGVORVQNEVKIKCOLKHPSTLE 74
DB 56 GAFPLKGGKAGKCFEISDPTKEVPFCKIVPKSLKPKHOKEMSMETSITHSLHSHOHVVG 115
QY 75 LYNVEDSNVYVYLVEKCHNGEMRNRYLKNRYKPFSENEARHPHQIITGMYLTHSHGIIH 134
DB 116 FHGFEDSDPFAVVYLETGRRSRILEIHKRR-KALTEPARAYYLQIVLGCQYLHNVYIH 174
QY 135 RDLTSLNLLTRNNMIKIALDFGLATQOLKMPHEKHYTLGGTNNYISPELATRSANGLESVD 194
DB 175 RDLKLGKGNFLNEDLEVKIGDFGLATKAYBGRKKTLCGTNNYLAPEVLAKGHSFEDV 234
QY 195 WSLGCMFTLLIGRPFDPTVTKTLANKVYLADYEMPSFLSEAKDLIHLRNPADRL 254
DB 235 WSLGIMVTLVGRKPFETSCLEKETYLRKNGEYSIPGHINPVASILQCKLQDPAPAR 294
QY 255 SLSSVLDHPFMSRNSSTKSLGIVEOSIDSGHATIS---TATYASSTISGSLFD-K 309
DB 295 TIHHLNDEFPT---SGYIPARLPITCLTIPPRFSIASSSLDPSN 336
QY 310 RRLI-----IGPLPNKMTVPFKPKSSSTDFFSSGSD-----NSFYQWGNQETS-N-SGR 357
DB 337 RKPFLTVANKGVENLPPDR---PREKEPVARETNEAIECHLSDLQQLTSVNAKPSER 392
QY 358 GRVYODAEERPH-----SRYLKRAVSPRSCTS-----NSGSAKTYTMR 398
DB 393 GLVROEAEAPACIPITVWSKWD---YSDKYGLGYOLCNSVGLFNDSTRILY----- 445
QY 399 CHSNAEMLSVKRSGCGENEREYPTDNNAN-----INPEFEKTS---SGSFERD 448
DB 446 -NNGDLSQYIERDQ---TESYLTVSHRPSLMKKTLLTLANFNNWSEHLKAGANITPR 500
QY 449 NNQALSNHLCPGKTPFPFADPTQETVQGF-----GNLQIN----- 486
DB 501 EGDELA-----RLPY-----LRTWERTRSAILIHLNSGTQVQINPQDHTKLIR 543
QY 487 AHUKTTEYDSISPNRDFQ 505
DB 544 GFLMAAVTY--INERKDFR 560

```

RESULT 12  
PLK1 HUMAN  
ID PLK1 HUMAN STANDARD; PRT; 603 AA.  
AC P53350;  
AD 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Serine/threonine-protein kinase PLK (EC 2.7.1.1-) (PLK-1) (Serine-threonine protein kinase 13) (STPK13).  
GN PLK OR PLK1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=94289293; PubMed=8018557;  
RA Hamanaka R., Malold S., Smith M.R., O'Connell C.D., Longo D.L., Ferris D.K.;  
RT "Cloning and characterization of human and murine homologues of the *Drosophila* polo serine-threonine kinase.";  
RL Cell Growth Differ. 5:249-257(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=94067140; PubMed=7902533;  
RA Lake R.J., Jelinek W.R.;  
RT "Cell cycle and terminal differentiation-associated regulation of the mouse mRNA encoding a conserved mitotic protein kinase.";  
RL Mol. Cell. Biol. 13:7793-7801(1993).  
RN [3]  
RP SEQUENCE FROM N.A.

```

RX MEDLINE=95051109; PubMed=7962193;  
RA Goleeeyn R.M., Schultz S.J., Bartek J., Ziemiecki A., Ried T., Nigg E.A.;  
RT "Cell cycle analysis and chromosomal localization of human Plk1, a putative homologue of the mitotic kinases Drosophila polo and Saccharomyces cerevisiae Cdc5.";  
RL J. Cell Sci. 107:1509-1517(1994).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=94173904; PubMed=8127874;  
RA Holtrich U., Wolf G., Braeuninger A., Karn T., Boehme B., Ruedemann-Waigmann H., Streibhardt K.;  
RT "Induction and down-regulation of PLK, a human serine/threonine kinase expressed in proliferating cells and tumors.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:1736-1740(1994).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon, and Lung;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F., Diatchenko L., Marnett K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fábry J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RP PHOSPHORYLATION OF SER-335.  
RX MEDLINE=22329537; PubMed=12442251;  
RA Wind M., Kelm O., Nigg E.A., Lehmann W.D.;  
RT "Identification of phosphorylation sites in the polo-like kinases Plk1 and Plk1 by a novel strategy based on element and electrospray high resolution mass spectrometry.";  
RL Proteomics 2:11516-11523(2002).  
CC -1- FUNCTION: May be required for cell division and may have a role during G1 or S phase.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- TISSUE SPECIFICITY: Placenta and colon.  
CC -1- DEVELOPMENTAL STAGE: Accumulates to a maximum during the G2 and M phases, declines to a nearly undetectable level following mitosis and throughout G1 phase, and then begins to accumulate again during S phase.  
CC -1- INDUCTION: By growth-stimulating agents.  
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
CC -1- SIMILARITY: CDC5/Polo subfamily.  
CC -1- SIMILARITY: Contains 2 POLO box domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial or send an email to license@isb-sib.ch).  
CC EMBL; U01038; AA56634.1; -;  
CC EMBL; I19559; AA36659.1; -;  
CC EMBL; X73458; CAA51837.1; -;  
CC EMBL; X75932; CAA53536.1; -;

```



```

DR EMBL: BC002369; AAH02369.1; -.
DR EMBL: BC003002; AAH03002.1; -.
DR EMBL: BC014846; AAH14846.1; -.
DR PIR: S34130; S34130.
DR Genew: HGNC:9077; PLK.
DR GK: P53350; -.
DR MIM: 602098; -.
DR GO: GO:0004674; P:protein serine/threonine kinase activity; TAS.
DR GO: GO:0008283; P:cell proliferation; TAS.
DR GO: GO:0007067; P:mitosis; TAS.
DR GO: GO:0007074; P:regulation of cell cycle; TAS.
DR InterPro: IPR000959; POLO box.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR InterPro: IPR002290; Ser_thr_kinase.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00659; POLO box; 2.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS0078; POLO BOX; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR Transferrase: Serine/threonine-protein kinase; ATP-binding; Repeat;
KW Nuclear protein; Phosphorylation.
FT DOMAIN 53 305
FT NP_BIND 59 67
FT BINDING 82 82
FT ACT_SITE 176 176
FT DOMAIN 417 480
FT DOMAIN 515 584
FT MOD_RES 335 335
FT CONFLICT 2 2
FT CONFLICT 11 11
FT CONFLICT 58 58
FT CONFLICT 60 60
FT CONFLICT 73 73
FT CONFLICT 141 141
FT CONFLICT 227 227
FT CONFLICT 301 301
FT CONFLICT 495 495
FT CONFLICT 501 501
SQ SEQUENCE 603 AA; 68254 MW; 178C2F31C10B8206 CRC64;

Query Match 10.6%; Score 536; DB 1; Length 603;
Best Local Similarity 32.2%; Pred. No. 2.3e-21;
Matches 118; Conservative 73; Mismatches 144; Indels 32; Gaps 6;

15 GNULGGRFAGVYRABSIHTGLEVAIKMIDKKMYAGWQRYQNEVKTHCOLKHSILE 74
56 GRFLGGRFAGVYRABSIHTGLEVAIKMIDKKMYAGWQRYQNEVKTHCOLKHSILE 115
75 LNYVEDSNVYVYVLEMGCHNGENRYYLKVYKPFSENEARHFMHQITITGMLYLSHGILH 134
116 FHFPEPDNPFVYVLEMGCHNGENRYYLKVYKPFSENEARHFMHQITITGMLYLSHGILH 174
135 RDLTSLNLLTRNNMTKIADFGIATQKPHKEGYLLCTGPNYISPEIATRSAGHSVDY 194
175 RDLKGNLFLNEDLEVKIGDFGLATKYVYGERKKTLCTGPNYIADBEVLSKKGHSFEVDY 234
195 WSLGCMFYLLTIGRPFDTDVNTKLVYLADEYMPSEFLSTAKDILHOLLARNPADRL 254
235 WSLGCMFYLLTIGRPFDTDVNTKLVYLADEYMPSEFLSTAKDILHOLLARNPADRL 294
255 SLSSVLDHFMNSSTKSDLTGVEDSIDSGHATIS--TATTAASSTISGSLPD-- 308
295 TINELENDSEPT-----SGYIPARLPITGLTIPRPSIAPSSLDPSN 336
309 -KRRLIIGQPLPKMYVFPKNSSTPSSSGDG-----NSFYQNGNQESN-SGGGRVY 361
337 RKPLTIVLNGLENPDRPERPREKEPVVRETGEVVDCHLSMDLQOLHSVNASKPSERGLVR 396
362 QDAEERP 368

```

```

DB 397 QDAEERP 403

RESULT 13
ID KKK1 YEAST STANDARD; PRT; 1518 AA.
AC P34214;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Probable serine/threonine-protein kinase YKL101W (BC 2.7.1.-).
GN YKL101W OR YKL453.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RX MEDLINE=94078677; PubMed=8256524;
RA Paller C., Valens M., Puzos V., Fukuhara H., Cheret G., Sor F.,
RT "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI
RT physically localizes the MBI gene and reveals eight new open reading
RT frames, including a homologue of the KIN1/KIN2 and SNF1 protein
RT kinases."
RL Yeast 9:1149-1155(1993).
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIM1
CC subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X71133; CAA50456.1; -.
DR EMBL: Z28101; CAA81941.1; -.
DR PIR: S37928; S37928.
DR HSSP: O63450; 1A06.
DR GenOnline: 139857; -.
DR SCD; S0001584; HSIL.
DR GO: GO:0005935; C:bud neck; IDA.
DR GO: GO:0005940; C:septin ring; IDA.
DR GO: GO:0004672; P:protein kinase activity; IDA.
DR GO: GO:0000086; P:G2/M transition of mitotic cell cycle; IGI.
DR GO: GO:0006489; P:protein amino acid phosphorylation; IDA.
DR GO: GO:0000074; P:regulation of cell cycle; IMP.
DR GO: GO:000135; P:septin checkpoint; IGI.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PD00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein; transferase; Serine/threonine-protein kinase;
KW ATP-binding.
FT DOMAIN 81 369
FT NP_BIND 87 95
FT BINDING 110 110
FT ACT_SITE 239 239
SQ SEQUENCE 1518 AA; 169592 MW; 803F84F7531241DD CRC64;

Query Match 10.4%; Score 526.5; DB 1; Length 1518;
Best Local Similarity 22.3%; Pred. No. 2.5e-20;

```



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).)

CC EMBL; X85758; CAA59766.1; -  
 CC EMBL; Z98559; CAB1167.1; -  
 DR PIR; T38254; T38254.  
 DR HSSP; O63450; 1A06.  
 DR GenBank; SPAC23C11.16; -  
 DR InterPro; IPR000959; Polo\_box.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00659; Polo\_box; 2.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PSS0078; POLO\_BOX; 2.  
 DR PROSITE; PSS00107; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PSS00108; PROTEIN\_KINASE\_ST; 1.  
 DR Transferrase; Serine/chreonine-protein kinase; ATP-binding; Repeat.  
 KW DOMAIN 41 236 PROTEIN KINASE.  
 FT NP\_BIND 47 55 ATP (BY SIMILARITY).  
 FT BINDING 69 69 ATP (BY SIMILARITY).  
 FT ACT\_SITE 163 163 BY SIMILARITY.  
 FT DOMAIN 500 567 POLO\_BOX\_1.  
 FT DOMAIN 604 670 POLO\_BOX\_2.  
 SQ SEQUENCE 683 AA; 77301 MW; F1ICD0EP98913917 CRC64;

Query Match 10.1%; Score 514.5; DB 1; Length 683;  
 Best Local Similarity 24.4%; Pred. No. 3.8e-20; Index 191; Gaps 22;  
 Matches 176; Conservative 106; Mismatches 247;

18 LGKSGFAGYVRAESHTGLVAIKMIDKAMVAGVORVONKVKIHCQLKPSILELYN 77  
 47 IEGEGFARCFRVVD-NYGNIVAKVAKSLQNDKTKLFGKIVKHQSHNITVGFID 105  
 78 YEDSNVYVLVLMCHNGENRKYKRVKPFSSNEARHFMHQITGMKLYHSHGILLRDL 137  
 106 CFEEDSTNIVLILIELCEHKSILMELLRRK-KOLTEPEVRYLMQITGLAKYMHKKRVVIRDL 164  
 138 TISNLLTNNMKIKIDGLAQLQKMPHEKHYTLCTPPVYISPEIA--TRSAHGESDW 195  
 165 KLGNIWLDSSNNVKGDFGLAALMDERKMTICGTPNTIAEILFNSKEGSPFVDLA 224  
 196 SLGCMFYTLILGSRPPEDTIVKNTLNKVVLADEYEMPS--FLSIEAKDLIHLRRPADR 253  
 225 SAGVWVYVALLIGKPPQDEKVKTIYRKIKANSYSPFSNDISAEADLISLLTHPSIR 284  
 254 ISLSVLDHPFMSRNSSTYSKDLGTVEDSIDSGHATITATVASSSTISGSLFDRRLI 313  
 285 PSIDIDVDFHEF-----FHTGYMAST----- 304  
 314 IGGPLENKTVTPKNSKST-----DESSSGDGSFYTQMG-----NOETSN--- 354  
 305 LDEEILHSPIMPWSSKSPQRIPLFVASASGVGNSAGVKNKPYALRTDEVNDRI 364  
 355 -----SGRGRVIO-----DAERPHSHVRLRAYSSDSGTSNS-----OSQA--- 391  
 365 LRSVLSPRRVRVPMKIGPRTKVPKSLSTALHAAKSTGDSIGSVKYLRESSQFVPT 424  
 392 KTYTWERCHSAEMLSVSKSGGGENEERYSPTDNNANIFNFFEKTSSTSSGSEPRPDNO 451  
 425 KSAVTQVPEPIQI:-----RSLSANVYRLSKVGNMKSDIWSYKATALKIG----- 471  
 452 ALSNHLCPKTPPRPADPPTQETVQVQWGNLQINHLKRTTEYDALS----- 499  
 472 -----MALSHHTALTSBEDADSEPVLFITTKVDY 500  
 500 PNRDFGHPDLQKDTSKNAWTDTKVKNKNSDASDNASVQOQT-----MKYNTALH 550

DB 501 SNKYGIGYOLSDSESVGHFNDDTSLFLFSADSEVEYALHPKDTIKPIYYPASKVESIR 560  
 QY SKPEIIQGRQCVGSDPLSRQSKTRGMEPPWGYNRTLSITSLVHRLKPIROKT----- 606  
 DB 561 SKIQLLHKFKSYWGONLSFAVODESEFKP--KNST--SNTLFWOHYLR-TRQAIMFRL 614  
 QY 607 -----KKAVVS-----ILDSEVECVELVKEVASQ-----EYVEVIO 638  
 DB 615 SNGIOPFNLRHKRVYISSTARKIIVLDKERREVELPLEGASAFSDLSRLKYNITETLE 674

RESULT 15  
 STK6\_HUMAN STANDARD; PRT; 403 AA.  
 AC 014965; 060445; 075873; 09B0D6; Q9UPG5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serine/chreonine kinase 6 (EC 2.7.1.37) (Serine/chreonine kinase 15)  
 DE (Aurora/Ipl1-related kinase 1) (Aurora-related kinase 1) (hARK1)  
 DE (Aurora-A) (Breast-tumor-amplified kinase).  
 GN STK6 OR STK15 OR Aik OR ARK1 OR AURA OR BTAk.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=97298083; PubMed=9153231;  
 RA Kimura M., Kotani S., Hattori T., Sumi N., Yoshioka T., Todokoro K., Okano Y.;  
 RT "Cell cycle-dependent expression and spindle pole localization of a novel human protein kinase, Aik, related to Aurora of Drosophila and yeast Ipl1.";  
 RT J. Biol. Chem. 272:13766-13771(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=98183439; PubMed=9514916;  
 RA Shindo M., Nakano H., Kuroyanagi H., Shiraawa T., Mihara M., Gilbert D.J., Jenkins N.A., Copeland N.G., Yagita H., Okumura K.;  
 RT "cDNA cloning, expression, subcellular localization, and chromosomal assignment of mammalian aurora homologues, aurora-related kinase (ARK) 1 and 2.";  
 RT 1 and 2.";  
 RL Biochem. Biophys. Res. Commun. 244:285-292(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast;  
 RX MEDLINE=98442657; PubMed=9771714;  
 RA Zhou H., Kuang J., Zhong L., Kuo W.-L., Gray J.W., Sahin A., Brinkley B.R., Sen S.;  
 RT "Tumour amplified kinase STK15/BTAk induces centrosome amplification, aneuploidy and transformation.";  
 RT Nat. Genet. 20:189-193(1998).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Wang L., Thibodeau S.N.;  
 RT "Functional analysis of the STK15 gene in human tumors.";  
 RT Submitted (OCT-1999) to the EMBL/GenBank/DBD databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.F., Babbage A.K., Bagguley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.B., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Grafham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E., Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lebraziaino M.H., Leversha M.A., Lloyd C., Lloyd D., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McComanche L.J., McMay K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillips B., J.C.T., Prathalingam S.R., Plumb R.W., Ramsey H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Walli M., Wallis J.M.,  
RA Whitehead S.L., Whitcaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming I., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.,  
RT "The DNA sequence and comparative analysis of human chromosome 20.",  
RL Nature 414:865-871(2001).  
RN [6].  
RP SEQUENCE FROM N.A.  
RC TISSUE=Cervix, Colon, Kidney, and Muscle;  
RX MEDLINE=22388257; PubMed=12471932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Cavaletto T.E.,  
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Cantini P., Prange C.,  
RA Rana S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gumarate P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Rahey J., Helton B., Ketteman M., Madan A., Rodriguez S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywiński M.J., Skalska U., Smalls D.E.,  
RA Schnerch A., Schein J., Krzywiński M.J., Skalska U., Smalls D.E.,  
RT "Generation and initial analysis of more than 15,000 full-length  
RL human and mouse cDNA sequences.",  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [7].  
RP CELL-CYCLE REGULATION.  
RX MEDLINE=21895866; PubMed=11790771;  
RA Tanaka M., Ueda A., Kanemori H., Ideguchi H., Yang J., Kitejima S.,  
RA Ishigatsubo Y.,  
RT "Cell-cycle-dependent regulation of human aurora A transcription is  
RL mediated by periodic repression of E2F1.",  
RN J. Biol. Chem. 277:10719-10726(2002).  
RN [8].  
RP REVIEW.  
RX MEDLINE=21306577; PubMed=11413462;  
RA Nigg E.A.;  
RT "Mitotic kinases as regulators of cell division and its checkpoints.",  
RL Nat. Rev. Mol. Cell Biol. 2:21-32(2001).  
CC -1- FUNCTION: May play a role in cell cycle regulation during anaphase  
and/or telophase, in relation to the function of the  
centrosome/spindle pole region during chromosome segregation.  
CC May be involved in microtubule formation and/or stabilization. May  
CC play a key role during tumor development and progression.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -1- SUBCELLULAR LOCATION: Localizes on centrosomes in interphase cells  
and at each spindle pole in mitosis.  
CC -1- TISSUE SPECIFICITY: Highly expressed in testis and weakly in  
skeletal muscle, thymus and spleen. Also highly expressed in  
colon, ovarian, prostate, neuroblastoma, breast and cervical  
cancer cell lines. Expression is cell-cycle regulated, low in  
G1/S, accumulates during G2/M, and decreases rapidly after.  
CC -1- PTM: Phosphorylated.  
CC -1- DISEASE: Defects in STK6 are responsible for numerical centrosome  
aberrations including aneuploidy.  
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.  
CC Aurora subfamily.  
CC -1- CAUTION: Ref.1 sequence differs from that shown due to frameshifts

Db	249	SKRVIRHDIKPELLELLGSAGELKIADFGMS--VHAPSSRRRTILCGTLDYLPPEMIEGRMH	306
Qy	189	GLESDVWSIGCMFYTLLIGRPFPDPTVKNLKVVLADYEMPSFLSTEAKDILHQLLR	248
Db	307	DERVDLWSLGVLCYEFLVGNKPPPEANTYQETVKRISRVEFTPPDPVTEGARDLISRLKH	366
Qy	249	NPADRLSLSSVLDHPFMSRNSSTKS	273
Db	367	NPSQRPMLKREVLHPWITANSKPS	391

Search completed: September 23, 2004, 21:00:51  
Job time : 18 secs

**This Page Blank (uspto)**





Query Match 99.9%; Score 5074; DB 4; Length 970;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 969; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATCI GEKIEDFKVGNLIGKGSFAGYRAESIHTGLEVAIKMIDKKAMTKAGMQRVONE 60  
 DB 1 MATCI GEKIEDFKVGNLIGKGSFAGYRAESIHTGLEVAIKMIDKKAMTKAGMQRVONE 60  
 QY 61 VKIHQOLKHPSTILELYNFEDSNYYVLVLEMCHNGEMNRYLKNRYKPSSEBARHFMHQI 120  
 DB 61 VKIHQOLKHPSTILELYNFEDSNYYVLVLEMCHNGEMNRYLKNRYKPSSEBARHFMHQI 120  
 QY 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHTYLCTGPNYISP 180  
 DB 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHTYLCTGPNYISP 180  
 QY 181 EIATRSAGLSDVWSLGCMPFTLLIGRPDDTDVKNLKNVVLADYEMPSFLSTEAKD 240  
 DB 181 EIATRSAGLSDVWSLGCMPFTLLIGRPDDTDVKNLKNVVLADYEMPSFLSTEAKD 240  
 QY 241 LIHOLLRRNPADRLSLSVLDHPFMSRNSSTSKDGLTVEDSIDSGHATITAITASSST 300  
 DB 241 LIHOLLRRNPADRLSLSVLDHPFMSRNSSTSKDGLTVEDSIDSGHATITAITASSST 300  
 QY 301 SISGLFDRRLILIGPLPNKMTVPFKNKSSTDFSSSGDGSFYTQMGNOETNSNGRGV 360  
 DB 301 SISGLFDRRLILIGPLPNKMTVPFKNKSSTDFSSSGDGSFYTQMGNOETNSNGRGV 360  
 QY 361 IODAEERPHSRYLRRAYSDRSGTSNSQOAKTYTMERCHSAEMLSVSKRSGGGENERY 420  
 DB 361 IODAEERPHSRYLRRAYSDRSGTSNSQOAKTYTMERCHSAEMLSVSKRSGGGENERY 420  
 QY 421 SPTDNNANFNFPFKETSSSGSFERPDNNQALSNLCPGKTPFPFADPTPOTETVQOAF 480  
 DB 421 SPTDNNANFNFPFKETSSSGSFERPDNNQALSNLCPGKTPFPFADPTPOTETVQOAF 480  
 QY 481 GNLQINAHLRKTTEDYSISPNRDFQGHPLQKDTSKNMTDTKVKKNKSDASDNASHVKOO 540  
 DB 481 GNLQINAHLRKTTEDYSISPNRDFQGHPLQKDTSKNMTDTKVKKNKSDASDNASHVKOO 540  
 QY 541 NNMKWTALHSHKEPITIQOECVFGSDPLSRQSKTRGMEPWCQONRTLSITPLVAHRLK 600  
 DB 541 NNMKWTALHSHKEPITIQOECVFGSDPLSRQSKTRGMEPWCQONRTLSITPLVAHRLK 600  
 QY 601 PIROKTKKAVVSIIDSEECVELVKEVYASQEVYEVLOISSDQNTITTYYPNGRGFPPLA 660  
 DB 601 PIROKTKKAVVSIIDSEECVELVKEVYASQEVYEVLOISSDQNTITTYYPNGRGFPPLA 660  
 QY 661 DRPSPPTDNIISRYSPDNLEPKYWKYQVARSFVQVLRSHSPKITTYFTRYAKCILMENSFG 720  
 DB 661 DRPSPPTDNIISRYSPDNLEPKYWKYQVARSFVQVLRSHSPKITTYFTRYAKCILMENSFG 720  
 QY 721 ADFEWFYVGVVGIHKTEDFIQVIEKTKGKSYTLKSBEENVSLKEELIKMYVDHANEGHRLCL 780  
 DB 721 ADFEWFYVGVVGIHKTEDFIQVIEKTKGKSYTLKSBEENVSLKEELIKMYVDHANEGHRLCL 780  
 QY 781 ALESIISSEERKTRSAFPFPIIIGRKPSTSPKALSPPSVDNYPTRDRASFNRMVWH 840  
 DB 781 ALESIISSEERKTRSAFPFPIIIGRKPSTSPKALSPPSVDNYPTRDRASFNRMVWH 840  
 QY 841 SAASPTQADILNPSVNTNFGILGTTTTASGTDIISNSLKCPLPKSAQLLKSVFKNVGNAT 900  
 DB 841 SAASPTQADILNPSVNTNFGILGTTTTASGTDIISNSLKCPLPKSAQLLKSVFKNVGNAT 900  
 QY 901 QLTSAVWVQFNDGSLVVOAGVSSISYTSPNGQTRVGENEKLDPYIKQKQCLSSILL 960  
 DB 901 QLTSAVWVQFNDGSLVVOAGVSSISYTSPNGQTRVGENEKLDPYIKQKQCLSSILL 960  
 QY 961 MFSNPTPNFH 970  
 DB 961 MFSNPTPNFH 970

RESULT 2  
 ID 000444 PRELIMINARY; PRT; 970 AA.  
 AC 000444;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Serine/threonine protein kinase.  
 GN SAK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA Karn T., Holtrich U., Wolf G., Hock B., Streibhardt K.,  
 RA Ruebsamen-Waigmann H.;  
 RT "Human SAK related to the PLK/polo family of cell cycle kinases shows  
 RT high mRNA expression in testis.";  
 RL Oncol. Rep. 4:505-510(1997).  
 DR EMBL: Y13115; CAA73575.1; -.  
 DR HSSP; Q63450; 1A06.  
 DR Genew; HGNC:11397; SYK18.  
 DR GO; GO:0004674; P:protein amino acid phosphorylation; TAS.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.  
 DR InterPro; IPR000559; POLO\_box.  
 DR InterPro; IPR000719; Prot\_Kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR Prodom; PD000001; Prot\_Kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS50078; POLO\_BOX; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE\_TYR; 1.  
 KW ATP-binding; kinase; Transferase.  
 SQ SEQUENCE 970 AA; 108955 MW; 552535615FBBE8DF CRC64;

Query Match 99.8%; Score 5067; DB 4; Length 970;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 967; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATCI GEKIEDFKVGNLIGKGSFAGYRAESIHTGLEVAIKMIDKKAMTKAGMQRVONE 60  
 DB 1 MATCI GEKIEDFKVGNLIGKGSFAGYRAESIHTGLEVAIKMIDKKAMTKAGMQRVONE 60  
 QY 61 VKIHQOLKHPSTILELYNFEDSNYYVLVLEMCHNGEMNRYLKNRYKPSSEBARHFMHQI 120  
 DB 61 VKIHQOLKHPSTILELYNFEDSNYYVLVLEMCHNGEMNRYLKNRYKPSSEBARHFMHQI 120  
 QY 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHTYLCTGPNYISP 180  
 DB 121 ITGMLYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHTYLCTGPNYISP 180  
 QY 181 EIATRSAGLSDVWSLGCMPFTLLIGRPDDTDVKNLKNVVLADYEMPSFLSTEAKD 240  
 DB 181 EIATRSAGLSDVWSLGCMPFTLLIGRPDDTDVKNLKNVVLADYEMPSFLSTEAKD 240  
 QY 241 LIHOLLRRNPADRLSLSVLDHPFMSRNSSTSKDGLTVEDSIDSGHATITAITASSST 300  
 DB 241 LIHOLLRRNPADRLSLSVLDHPFMSRNSSTSKDGLTVEDSIDSGHATITAITASSST 300  
 QY 301 SISGLFDRRLILIGPLPNKMTVPFKNKSSTDFSSSGDGSFYTQMGNOETNSNGRGV 360  
 DB 301 SISGLFDRRLILIGPLPNKMTVPFKNKSSTDFSSSGDGSFYTQMGNOETNSNGRGV 360  
 QY 361 IODAEERPHSRYLRRAYSDRSGTSNSQOAKTYTMERCHSAEMLSVSKRSGGGENERY 420  
 DB 361 IODAEERPHSRYLRRAYSDRSGTSNSQOAKTYTMERCHSAEMLSVSKRSGGGENERY 420

```

Db 361 IODAEERPHSRYLRAVSSDRSGTSNSQSOAKTYTWERCHSAEMLSVSKRSGGGENEERY 420
Qy 421 SPTDNNANIFNPFKEKTSSSSGSFERRDNNQALSNHLCGKTPFPADPTPOTETVQOWF 480
Db 421 SPTDNNANIFNPFKEKTSSSSGSFERRDNNQALSNHLCGKTPFPADPTPOTETVQOWF 480
Qy 481 GNLIQINAHARKTTEYDVISPNRDFQHPDLQKOTSQNAATDTYVKKQSDADNAHVSXQO 540
Db 481 GNLIQINAHARKTTEYDVISPNRDFQHPDLQKOTSQNAATDTYVKKQSDADNAHVSXQO 540
Qy 541 NTMKWTALHSKPEIIQOECVFCSDPLSEOSKTRGMEPPMGVQNTLRSTSTPLVAHRLK 600
Db 541 NTMKWTALHSKPEIIQOECVFCSDPLSEOSKTRGMEPPMGVQNTLRSTSTPLVAHRLK 600
Qy 601 PIROKTKKAVVSIIDSEEVCELVKEVYASQEVYKEVLQISSDGNITITIIYPNGRGFPPLA 660
Db 601 PIROKTKKAVVSIIDSEEVCELVKEVYASQEVYKEVLQISSDGNITITIIYPNGRGFPPLA 660
Qy 661 DRPPSPTDNISRYSPNLPBKTWRKYQVARSFVQVRSKPKITTYTRAKCLIMENSPG 720
Db 661 DRPPSPTDNISRYSPNLPBKTWRKYQVARSFVQVRSKPKITTYTRAKCLIMENSPG 720
Qy 721 ADFEWFYDGVKIHKTEDFIYIEKTKSYTLKSEBVSNSLKEEIMYMDHANEGHRI 780
Db 721 ADFEWFYDGVKIHKTEDFIYIEKTKSYTLKSEBVSNSLKEEIMYMDHANEGHRI 780
Qy 781 ALESIISEBERKTRAPFPPIIIGKRPSTSPKALSPPPSVDSNTPTDRASFNMVNM 840
Db 781 ALESIISEBERKTRAPFPPIIIGKRPSTSPKALSPPPSVDSNTPTDRASFNMVNM 840
Qy 841 SAASPTQALILPMSWNTNGLGLTTASGTDISSNLKCLPKSAQLLKSVPKANGMAT 900
Db 841 SAASPTQALILPMSWNTNGLGLTTASGTDISSNLKCLPKSAQLLKSVPKANGMAT 900
Qy 901 QLTSGAVWQFNDGSQLVYQAGVSSISYTSPPNGQTRRGENKLPYIKQKQLCSILL 960
Db 901 QLTSGAVWQFNDGSQLVYQAGVSSISYTSPPNGQTRRGENKLPYIKQKQLCSILL 960
Qy 961 MFSNPTPNFH 970
Db 961 MFSNPTPNFH 970

```

```

DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser thr kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam: PF00069; kinase_1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR PROSITE: PS00107; POLO_BOX; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR ATP-binding; Transferase.
SQ SEQUENCE 970 AA; 109084 MW; 1F431PA607A3550A CRC64;

Query Match 99.6%; Score 5057; DB 4; Length 970;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 965; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

1 MATCI GEKIEDPFGVNLGKSGPAGYVRAESIHTGLEVAIKMIDKKAMTKAGVQGVONE 60
1 MATCI GEKIEDPFGVNLGKSGPAGYVRAESIHTGLEVAIKMIDKKAMTKAGVQGVONE 60
61 VKIHQKLPSTILELYNPFEDSNYYVLVLEMGCHGEMNRYLKNRYVPSSENEARHFMQI 120
61 VKIHQKLPSTILELYNPFEDSNYYVLVLEMGCHGEMNRYLKNRYVPSSENEARHFMQI 120
121 ITGMLYLSHGILHRDLTSLNLLTRNNMIKIADFGLATQKMPHEKHYTLGCTPYIIS 180
121 ITGMLYLSHGILHRDLTSLNLLTRNNMIKIADFGLATQKMPHEKHYTLGCTPYIIS 180
181 EIAATSAHGLSDVMSIGCMFTLLIGRPFPDVTYKNTLKNVVLADYEMPSFLSEAKD 240
181 EIAATSAHGLSDVMSIGCMFTLLIGRPFPDVTYKNTLKNVVLADYEMPSFLSEAKD 240
241 LIHQLLRRPADRLSLSSVLDHPFMSRNSSTKSLGTVEDSIDSGHATISTAITASSST 300
241 LIHQLLRRPADRLSLSSVLDHPFMSRNSSTKSLGTVEDSIDSGHATISTAITASSST 300
301 S1SGSLFDRKRLILGQPLFNKNTVFPKAKSSIDFSSSGDGNSTYTWQNGOEFNSGRGV 360
301 S1SGSLFDRKRLILGQPLFNKNTVFPKAKSSIDFSSSGDGNSTYTWQNGOEFNSGRGV 360
361 IODAEERPHSRYLRAVSSDRSGTSNSQSOAKTYTWERCHSAEMLSVSKRSGGGENEERY 420
361 IODAEERPHSRYLRAVSSDRSGTSNSQSOAKTYTWERCHSAEMLSVSKRSGGGENEERY 420
421 SPTDNNANIFNPFKEKTSSSSGSFERRDNNQALSNHLCGKTPFPADPTPOTETVQOWF 480
421 SPTDNNANIFNPFKEKTSSSSGSFERRDNNQALSNHLCGKTPFPADPTPOTETVQOWF 480
481 GNLIQINAHARKTTEYDVISPNRDFQHPDLQKOTSQNAATDTYVKKQSDADNAHVSXQO 540
481 GNLIQINAHARKTTEYDVISPNRDFQHPDLQKOTSQNAATDTYVKKQSDADNAHVSXQO 540
541 NTMKWTALHSKPEIIQOECVFCSDPLSEOSKTRGMEPPMGVQNTLRSTSTPLVAHRLK 600
541 NTMKWTALHSKPEIIQOECVFCSDPLSEOSKTRGMEPPMGVQNTLRSTSTPLVAHRLK 600
601 PIROKTKKAVVSIIDSEEVCELVKEVYASQEVYKEVLQISSDGNITITIIYPNGRGFPPLA 660
601 PIROKTKKAVVSIIDSEEVCELVKEVYASQEVYKEVLQISSDGNITITIIYPNGRGFPPLA 660
661 DRPPSPTDNISRYSPNLPBKTWRKYQVARSFVQVRSKPKITTYTRAKCLIMENSPG 720
661 DRPPSPTDNISRYSPNLPBKTWRKYQVARSFVQVRSKPKITTYTRAKCLIMENSPG 720
721 ADFEWFYDGVKIHKTEDFIYIEKTKSYTLKSEBVSNSLKEEIMYMDHANEGHRI 780
721 ADFEWFYDGVKIHKTEDFIYIEKTKSYTLKSEBVSNSLKEEIMYMDHANEGHRI 780
781 ALESIISEBERKTRAPFPPIIIGKRPSTSPKALSPPPSVDSNTPTDRASFNMVNM 840
781 ALESIISEBERKTRAPFPPIIIGKRPSTSPKALSPPPSVDSNTPTDRASFNMVNM 840

```

Db 781 ALBSIISEERKTRAPFPPIIIGRKPOSTSSPKALSPPSVDSNPTPRDASFNRMVMH 840  
Qy 841 SAASPTQAPILNPMWMTNEGILTTTASGMDISSNSLKDCIPKSAQLLKSVFNVMAT 900  
Db 841 SDASPTQAPILNPMWMTNEGILTTTASGMDISSNSLKDCIPKSAQLLKSVFNVMAT 900  
Qy 901 QLTSGAVWVQNDQSOLVVOAGVSSISYTSNPGQTTTNGENEKLPDYIKOKLQCLSSIL 960  
Db 901 QLTSGAVWVQNDQSOLVVOAGVSSISYTSNPGQTTTNGENEKLPDYIKOKLQCLSSIL 960  
Qy 961 MFSNPTNPFH 970  
Db 961 MFSNPTNPFH 970

RESULT 4  
Q64702 PRELIMINARY; PRT; 925 AA.  
ID 064702;  
AC 064702;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE SMK/PLK-AKIN kinase (Protein kinase SMK/PLK-AKIN) (EC 2.7.1.1-).  
GN STK18 OR SAK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DRA/2; TISSUE=Lymphoma;  
RX MEDLINE=94294387; PubMed=8022793;  
RA Fode C., Motro B., Yousefi S., Heffernan M., Dennis J.W.;  
RT "Sak, a murine protein-serine/threonine kinase that is related to the  
RT Drosophila polo kinase and involved in cell proliferation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:6388-6392(1994).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Sak-a;  
CC IsoId=Q64702-1; Sequence=Displayed;  
CC Name=Sak-b;  
CC IsoId=Q64702-2; Sequence=VSP\_050448, VSP\_050449;  
CC EMBL, I29479; AAC37648.1; -;  
DR EMBL, I29480; AAC37649.1; -;  
DR PIR, A55748; A55748.  
DR HSSP, Q00534; IBI8.  
DR MGD, MGI:101783; Stk18.  
DR GO, GO:0005524; F:ATP binding; IEA.  
DR GO, GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO, GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO, GO:0016740; F:transferase activity; IEA.  
DR GO, GO:0007049; F:cell cycle; IEA.  
DR GO, GO:0006468; F:protein amino acid phosphorylation; IEA.  
DR InterPro, IPR000959; POLO box.  
DR InterPro, IPR000719; ProC kinase.  
DR InterPro, IPR002290; Ser Thr kinase.  
DR InterPro, IPR008266; Tyr Thr kinase.  
DR Pfam, PF00069; Kinase; 1.  
DR ProDom, PD000001; ProC kinase; 1.  
DR SMART, SM00220; S\_TKc; 1.  
DR PROSITE, PS00107; POLO BOX; 1.  
DR PROSITE, PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE, PS00011; PROTEIN KINASE DOM; 1.  
DR PROSITE, PS00109; PROTEIN KINASE\_TYR; 1.  
KM ATP-binding; Alternative protein; Serine/threonine-protein kinase;  
KM transferase.  
FT TRANSPLIC 417 464  
FT SNHCLGKTPFPFADOTPEMVMQVQGNLOMNAHGETM  
FT EHHVTSP -> RYSPTRKSNVAVLTSINRKOPIYVDLKKRI  
FT MTEQKMDLNLINKFDR (in isoform Sak-b).  
FT VARSPLIC 465 925  
FT Missing (in isoform Sak-b).  
FT Missing (in isoform Sak-b).  
FT Missing (in isoform Sak-b).  
SQ SEQUENCE 925 AA; 103685 MW; D868A76BB7343E81 CRC64;

Query Match 77.3%; Score 3927.5; DB 11; Length 925;  
Best Local Similarity 78.6%; Pred. No. 2.2e-248;  
Matches 763; Conservative 76; Mismatches 83; Indels 49; Gaps 9;  
Qy 1 MATCGEKIEDPKVGNLKGSPAGYVRAESIHGTEVAIKIIDKKAYKAGMVRVONE 60  
Db 1 MACIGERLEDPKVGALLGKSPAGYVRAESIHGTEVAIKIIDKKAYKAGMVRVONE 60  
Qy 61 VKIHQQLKHPSTLELYNYPEDSNYYVYLVMCHNEMNRYLKNRYKPSSEBARHPMHOI 120  
Db 61 VKIHQQLKHPSTLELYNYPEDSNYYVYLVMCHNEMNRYLKNRYKPSSEBARHPMHOI 120  
Qy 121 ITGMLYHSHGILHMDLTLSNLLTRNNYIKADFGALATOLKMPHEKHYTLGTPNYTSP 180  
Db 121 ITGMLYHSHGILHMDLTLSNLLTRNNYIKADFGALATOLKMPHEKHYTLGTPNYTSP 180  
Qy 181 E1ATRSAGHLESDVMSLCMPFYTLIGRPPTDTPVKNLTKVYLADVEMSPLEAKD 240  
Db 181 E1ATRSAGHLESDVMSLCMPFYTLIGRPPTDTPVKNLTKVYLADVEMSPLEAKD 240  
Qy 241 LIHQLLRNPADRLSLSSVLDHPFMSRNSSTKSKDGLVEDSDISGHATISTATTAASST 300  
Db 241 LIHQLLRNPADRLSLSSVLDHPFMSRNSSTKSKDGLVEDSDISGHATISTATTAASST 300  
Qy 301 SISGSLPDKRLLIGQPLPKKATVPKKKSTDPSSGDSGSPYTOGN--QETSNSGRG 358  
Db 301 SISGSLD-RRLLVGQPLPKKATVPKKKSTDPSSGDSGSPYTOGN--QETSNSGRG 358  
Qy 301 SLSSGLD-RRLLVGQPLPKKATVPKKKSTDPSSGDSGSPYTOGN--QETSNSGRG 358  
Db 301 SLSSGLD-RRLLVGQPLPKKATVPKKKSTDPSSGDSGSPYTOGN--QETSNSGRG 358  
Qy 359 RVIQDAEERPHSRYLRAVSSDRSGTSNSQAKTYTWERCHSABMVSYSRSGGENEE 418  
Db 359 RVIQDAEERPHSRYLRAVSSDRSGTSNSQAKTYTWERCHSABMVSYSRSGGENEE 418  
Qy 359 RVIQDAEERPHSRYLRAVSSDRSGTSNSQAKTYTWERCHSABMVSYSRSGGENEE 418  
Db 359 RVIQDAEERPHSRYLRAVSSDRSGTSNSQAKTYTWERCHSABMVSYSRSGGENEE 418  
Qy 419 RYSPIDNNANIFNPFKEKTSSTSSGSEFEPDNNQALSNHLCPGKTPPPADTPQETVQ 478  
Db 419 RYSPIDNNANIFNPFKEKTSSTSSGSEFEPDNNQALSNHLCPGKTPPPADTPQETVQ 478  
Qy 411 -----LDENQSSNHCLGKTPFPFADTPQEMVQ 442  
Db 411 -----LDENQSSNHCLGKTPFPFADTPQEMVQ 442  
Qy 479 WFGNLOINAHLRKTXEYDSISPNRDPQGHPLQKDTSKMAMTQKVKNSDASINMAYK 538  
Db 479 WFGNLOINAHLRKTXEYDSISPNRDPQGHPLQKDTSKMAMTQKVKNSDASINMAYK 538  
Qy 443 WFGNLOINAHLRKTXEYDSISPNRDPQGHPLQKDTSKMAMTQKVKNSDASINMAYK 501  
Db 443 WFGNLOINAHLRKTXEYDSISPNRDPQGHPLQKDTSKMAMTQKVKNSDASINMAYK 501  
Qy 539 QONTMKTMTALSHKPBILIOECVFSGLSEOSKTRGMEPPWGYONTLSITSPLVAHR 598  
Db 539 QONTMKTMTALSHKPBILIOECVFSGLSEOSKTRGMEPPWGYONTLSITSPLVAHR 598  
Qy 502 QLSAMKRYSAHHRKPEVMPQEP--GLPHSEQSKNRMESITLQYKQKTLASISPLAHR 559  
Db 502 QLSAMKRYSAHHRKPEVMPQEP--GLPHSEQSKNRMESITLQYKQKTLASISPLAHR 559  
Qy 599 LKPIROKTKKAVVSIIDSEVCELVKEXASQRYKXEVLOISSDQNTITTYPNGRGFP 658  
Db 599 LKPIROKTKKAVVSIIDSEVCELVKEXASQRYKXEVLOISSDQNTITTYPNGRGFP 658  
Qy 560 LKPIROKTKKAVVSIIDSEVCELVKEXASQRYKXEVLOISSDQNTITTYPNGRGFP 619  
Db 560 LKPIROKTKKAVVSIIDSEVCELVKEXASQRYKXEVLOISSDQNTITTYPNGRGFP 619  
Qy 659 LADRPSPPTDNIISRYSPDNLPKRYWRYQYASRFVOLVRSKSPKITTYFTFYACILMENS 718  
Db 659 LADRPSPPTDNIISRYSPDNLPKRYWRYQYASRFVOLVRSKSPKITTYFTFYACILMENS 718  
Qy 620 LADRPSPPTDNIISRYSPDNLPKRYWRYQYASRFVOLVRSKSPKITTYFTFYACILMENS 679  
Db 620 LADRPSPPTDNIISRYSPDNLPKRYWRYQYASRFVOLVRSKSPKITTYFTFYACILMENS 679  
Qy 719 PGADFEVWFDGVKIKHTEDFIOVIEKTKGSYTLKSESEVNSLKEIKMYMDHANEGRRI 778  
Db 719 PGADFEVWFDGVKIKHTEDFIOVIEKTKGSYTLKSESEVNSLKEIKMYMDHANEGRRI 778  
Qy 680 PGADFEVWFDGVKIKHTEDFIOVIEKTKGSYTLKSESEVNSLKEIKMYMDHANEGRRI 729  
Db 680 PGADFEVWFDGVKIKHTEDFIOVIEKTKGSYTLKSESEVNSLKEIKMYMDHANEGRRI 729  
Qy 779 CLAESIISEERKTRAPFPPIIIGRKPGSTSPKALSPPSVDSNPTPRDASFNRMVMH 838  
Db 779 CLAESIISEERKTRAPFPPIIIGRKPGSTSPKALSPPSVDSNPTPRDASFNRMVMH 838  
Qy 740 CLAESIISEERKTRAPFPPIIIGRKPGSTSPKALSPPSVDSNPTPRDASFNRMVMH 798  
Db 740 CLAESIISEERKTRAPFPPIIIGRKPGSTSPKALSPPSVDSNPTPRDASFNRMVMH 798  
Qy 839 MHSASPTQAPILNPMWMTNEGILTTTASGMDISSNSLKDCIPKSAQLLKSVFNVMAT 898  
Db 839 MHSASPTQAPILNPMWMTNEGILTTTASGMDISSNSLKDCIPKSAQLLKSVFNVMAT 898  
Qy 799 VNSAAPTQSPGSLPSTIVTEGILGHATNATGTVSS-----LPKSAQLLKSVFNVMAT 853  
Db 799 VNSAAPTQSPGSLPSTIVTEGILGHATNATGTVSS-----LPKSAQLLKSVFNVMAT 853  
Qy 899 ATQLTSGAVWVQNDQSOLVVOAGVSSISYTSNPGQTTTNGENEKLPDYIKOKLQCLSSI 958  
Db 899 ATQLTSGAVWVQNDQSOLVVOAGVSSISYTSNPGQTTTNGENEKLPDYIKOKLQCLSSI 958  
Qy 854 ATQLTSGAVWVQNDQSOLVVOAGVSSISYTSNPGQTTTNGENEKLPDYIKOKLQCLSSI 913  
Db 854 ATQLTSGAVWVQNDQSOLVVOAGVSSISYTSNPGQTTTNGENEKLPDYIKOKLQCLSSI 913  
Qy 959 LMFNSPTNPFH 969  
Db 959 LMFNSPTNPFH 969  
Qy 914 LMFNSPTNPFH 924  
Db 914 LMFNSPTNPFH 924

```

RESULT 5
Q8R015 PRELIMINARY; PRT; 925 AA.
ID Q8R015;
AC Q8R015;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN STK18.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026785; AAH26785.1; -.
DR PDB; 1MBY; 30-OCT-02.
DR MGI; MGI:101783; Stk18.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0007049; P:cell cycle; IEA.
DR GO; GO:0004468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000959; POLO_box.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PSS0078; POLO_BOX; 1.
DR PROSITE; PSS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PSS00109; PROTEIN_KINASE_TYR; 1.
DR Hypothetical protein; ATP-binding; Transferase.
KW SEQUENCE 925 AA; 103861 MW; 844AF8C9A5C4C1 CRC64;

Query Match 77.2%; Score 3918.5; DB 11; Length 925;
Best Local Similarity 78.4%; Pred. No. 8.6e-248;
Matches 761; Conservative 77; Mismatches 84; Indels 49; Gaps 9;

QY 1 MATCIKEKIDPFVGNLKGKSPAGYVRAISHTGLGVAIKMIKKAMYGAVQVRQNS 60
DB 1 MAACIGERIEDPRVGNLKGKSPAGYVRAISHTGLGVAIKMIKKAMYGAVQVRQNS 60
QY 61 VKIHQQLKHPSTILELYNFEDSNYYVLVLEMCNGEMNRYLKNRVKPSSENAHFHQI 120
DB 61 VKIHQQLKHPSTILELYNFEDSNYYVLVLEMCNGEMNRYLKNRVKPSSENAHFHQI 120
QY 61 VKIHQQLKHPSTILELYNFEDSNYYVLVLEMCNGEMNRYLKNRVKPSSENAHFHQI 120
DB 61 VKIHQQLKHPSTILELYNFEDSNYYVLVLEMCNGEMNRYLKNRVKPSSENAHFHQI 120
QY 121 ITGMLYLHSHGILHRLTISNLLTRNNIKIADFGLAQLKMPHKHTYLCTPRYIS 180
DB 121 ITGMLYLHSHGILHRLTISNLLTRNNIKIADFGLAQLKMPHKHTYLCTPRYIS 180
QY 121 ITGMLYLHSHGILHRLTISNLLTRNNIKIADFGLAQLKMPHKHTYLCTPRYIS 180
DB 121 ITGMLYLHSHGILHRLTISNLLTRNNIKIADFGLAQLKMPHKHTYLCTPRYIS 180
QY 181 EIAIRSAHGLSDVMSLGCMFYLLIGRPFDTDVYNTLANKVLDVEMPSFLISLEAK 240
DB 181 EIAIRSAHGLSDVMSLGCMFYLLIGRPFDTDVYNTLANKVLDVEMPSFLISLEAK 240
QY 181 EIAIRSAHGLSDVMSLGCMFYLLIGRPFDTDVYNTLANKVLDVEMPSFLISLEAK 240
DB 181 EIAIRSAHGLSDVMSLGCMFYLLIGRPFDTDVYNTLANKVLDVEMPSFLISLEAK 240
QY 241 LIHQLLRRNPADRLSSVLDHPFMSRNSSTKSKDLGVEDSIDSGHATISTAITASSST 300
DB 241 LIHQLLRRNPADRLSSVLDHPFMSRNSSTKSKDLGVEDSIDSGHATISTAITASSST 300
QY 241 LIHQLLRRNPADRLSSVLDHPFMSRNSSTKSKDLGVEDSIDSGHATISTAITASSST 300
DB 241 LIHQLLRRNPADRLSSVLDHPFMSRNSSTKSKDLGVEDSIDSGHATISTAITASSST 300
QY 301 STSGSLFDRBRILLIGPLPKMKVFPKNSSTDFSSGDSNCTGMGN--OETNSGRG 358
DB 301 STSGSLFDRBRILLIGPLPKMKVFPKNSSTDFSSGDSNCTGMGN--OETNSGRG 358
QY 301 STSGSLFDRBRILLIGPLPKMKVFPKNSSTDFSSGDSNCTGMGN--OETNSGRG 358
DB 301 STSGSLFDRBRILLIGPLPKMKVFPKNSSTDFSSGDSNCTGMGN--OETNSGRG 358
QY 359 RVYDAEERPHSRYLRAVSDRSGTNSGSAKTYTMEKCHSAEMLSVSKRSGGENEB 418
DB 359 RVYDAEERPHSRYLRAVSDRSGTNSGSAKTYTMEKCHSAEMLSVSKRSGGENEB 418

```

```

DB 359 RVYDAEERPHSRYLRAVSDRSGTNSGSAKTYTMEKCHSAEMLSVSKRSGGENEB 418
QY 419 RYSPIDNNANINFPREKTSSSSGSERPDNNQALSNHLCPKTPFPADPTPQETVQO 478
DB 419 RYSPIDNNANINFPREKTSSSSGSERPDNNQALSNHLCPKTPFPADPTPQETVQO 478
QY 411 -----LDENQSSNNHCLGKTPFPADPTPQETVQO 442
DB 411 -----LDENQSSNNHCLGKTPFPADPTPQETVQO 442
QY 479 WFGNLIQINAHLRKTEYDISPNRDFOGHPDLQKDTSKNAWTDYKKNKSDSDNASHYK 538
DB 479 WFGNLIQINAHLRKTEYDISPNRDFOGHPDLQKDTSKNAWTDYKKNKSDSDNASHYK 538
QY 443 WFGNLIQINAHLRKTEYDISPNRDFOGHPDLQKDTSKNAWTDYKKNKSDSDNASHYK 501
DB 443 WFGNLIQINAHLRKTEYDISPNRDFOGHPDLQKDTSKNAWTDYKKNKSDSDNASHYK 501
QY 539 QONTMKTATLHSPKEITIOECVFGSDPLSEQSKTRGMEPMGYONRTLRSTSPVANH 598
DB 539 QONTMKTATLHSPKEITIOECVFGSDPLSEQSKTRGMEPMGYONRTLRSTSPVANH 598
QY 502 QLSAMKYNASHHKKEVMPQEP--DLHPHSEOSKRSKMSSTLGYQKPTLRSTSPVANH 559
DB 502 QLSAMKYNASHHKKEVMPQEP--DLHPHSEOSKRSKMSSTLGYQKPTLRSTSPVANH 559
QY 599 LKPIRQKTKKAVSLIDSEVCVELKESVAGSEVYVQISDGGTITITYPNGRGPR 658
DB 599 LKPIRQKTKKAVSLIDSEVCVELKESVAGSEVYVQISDGGTITITYPNGRGPR 658
QY 560 LKPIRQKTKKAVSLIDSEVCVELKESVAGSEVYVQISDGGTITITYPNGRGPR 619
DB 560 LKPIRQKTKKAVSLIDSEVCVELKESVAGSEVYVQISDGGTITITYPNGRGPR 619
QY 659 LADRPSPPTDNISRYSPFNLPEKYWRKYQYASRFVQVYKSKPKITYFTRVAKCILMENS 718
DB 659 LADRPSPPTDNISRYSPFNLPEKYWRKYQYASRFVQVYKSKPKITYFTRVAKCILMENS 718
QY 620 LADRPSPPTDNISRYSPFNLPEKYWRKYQYASRFVQVYKSKPKITYFTRVAKCILMENS 679
DB 620 LADRPSPPTDNISRYSPFNLPEKYWRKYQYASRFVQVYKSKPKITYFTRVAKCILMENS 679
QY 719 PADPEFWYFDGKIKHTEDEPIQVIEKTKSYTLKSESFVNSLKEEIKMYMDHANGHRI 778
DB 719 PADPEFWYFDGKIKHTEDEPIQVIEKTKSYTLKSESFVNSLKEEIKMYMDHANGHRI 778
QY 680 PADPEFWYFDGKIKHTEDEPIQVIEKTKSYTLKSESFVNSLKEEIKMYMDHANGHRI 739
DB 680 PADPEFWYFDGKIKHTEDEPIQVIEKTKSYTLKSESFVNSLKEEIKMYMDHANGHRI 739
QY 779 CLATESIIEERKTRSAFPFIIIGRKGRGSSPKALSPPSVDSNYTRDRASFNRY 838
DB 779 CLATESIIEERKTRSAFPFIIIGRKGRGSSPKALSPPSVDSNYTRDRASFNRY 838
QY 740 CILBSVISEEERKTRSAFPFIIIGRKGRGSSPKALSPPSVDSNYTRDRASFNRY 798
DB 740 CILBSVISEEERKTRSAFPFIIIGRKGRGSSPKALSPPSVDSNYTRDRASFNRY 798
QY 839 MHSASPTQAPILNPSMVTNEGILGTTTASGTDISGNSLKDCLPKSAQLLKSFVQVQGM 898
DB 839 MHSASPTQAPILNPSMVTNEGILGTTTASGTDISGNSLKDCLPKSAQLLKSFVQVQGM 898
QY 799 VNSAAPPTQAPILNPSMVTNEGILGTTTASGTDISGNSLKDCLPKSAQLLKSFVQVQGM 853
DB 799 VNSAAPPTQAPILNPSMVTNEGILGTTTASGTDISGNSLKDCLPKSAQLLKSFVQVQGM 853
QY 899 ATQITSGAVVQFNDGSLVQAGVSSISYSPNGQTRTYGNEKLPDYIKQQLCLSSI 958
DB 899 ATQITSGAVVQFNDGSLVQAGVSSISYSPNGQTRTYGNEKLPDYIKQQLCLSSI 958
QY 854 ATQITSGAVVQFNDGSLVQAGVSSISYSPNGQTRTYGNEKLPDYIKQQLCLSSI 913
DB 854 ATQITSGAVVQFNDGSLVQAGVSSISYSPNGQTRTYGNEKLPDYIKQQLCLSSI 913
QY 959 LMFSPNTPNF 969
DB 959 LMFSPNTPNF 924

RESULT 6
Q9CVU6 PRELIMINARY; PRT; 535 AA.
ID Q9CVU6;
AC Q9CVU6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Serine/threonine kinase 18 (Fragment).
GN STK18.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Testis;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai T., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Kadoya K., Matsumoto H.A., Ashburner M., Batalov S., Casavant T.,
RA Pleischmann W., Gaasterland T., Glasl C., King B., Kochiwa H.,
RA Kuenli P., Lewis S., Matsuo Y., Nishikido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Oikio T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.F.,
RA Gustinchik S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

```

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyszew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.",  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK006459; BAB24599.1; -.  
 DR MGD; MGI:101783; Stk18.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0007049; P:cell cycle; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000959; POLO box.  
 DR PROSITE; PS50078; POLO\_Box; 1.  
 FT NON TER 1  
 SQ SEQUENCE 535 AA; 59788 MW; 3A35E0713FD641B3 CRC64;  
 Query Match 41.7%; Score 2116; DB 11; Length 535;  
 Best Local Similarity 71.6%; Pred. No. 4e-130;  
 Matches 414; Conservative 53; Mismatches 67; Indels 44; Gaps 5;  
 QY 392 KTYTMRCHSAEMLSVSKRGSGGSENERISPTNNANITFPEKETSSSGSFEPDNNQ 451  
 DB 1 KTSVSECHSVEMLSKRRS-----LDENQ 25  
 QY 452 ALSNHLCPGKTPFPADPTPQTEVVOFGNLIQINHLKRTTEYDTSIPNPDQGHDD 511  
 DB HSNHHCIGKTPFPADPTPQTEVVOFGNLIQINHLKRTTEYDTSIPNPDQGHDD 85  
 QY 512 KOTSKAATDTKAKKSDASDANSHYKQONTKMTALSKREIIQOECVFGSDPLSEOS 571  
 DB 86 -DILRNAWDTTRASKADTSANVAHYKQLSAMKYVAHKKHKEVPMQEP--GLHPHSEGS 142  
 QY 572 KTRGMPWQYQRTLRSTSPVLRKPIRQKTKKAVSLIDSEVCELVKYEASG 631  
 DB 143 KRNSESTIGYQPTLRSTSPVLRKPIRQKTKKAVSLIDSEVCELVKYEASG 202  
 QY 632 YKAEVLIQSSDGTITTYPNGRGFPPLADRPSPPTDNISSRYSDNLPEKRYKQYASR 691  
 DB 203 YKAEVLIQSSDGTITTYPNGRGFPPLADRPSPPTDNISSRYSDNLPEKRYKQYASR 262  
 QY 692 PVLVSKSPKITYFRYAKCLIMENSPEADPEWVYDVKTHKEDFQVTEKTSYT 751  
 DB 263 FVLVSKSPKITYFRYAKCLIMENSPEADPEWVYDVKTHKEDFQVTEKTSYT 322  
 QY 752 LKSESEVNSLKEEIKMYMDHANEGRICLSEIISSEERKTRSAFPFPIIIGKPGSTS 811  
 DB 323 LKNEVEVTLKEEVKMYMDHANEGRICLSEIISSEERKTRSAFPFPIIIGKPGSTS 382  
 QY 812 SPKALSPSPSVUNPSTRDASFNRMVMSAASPQADPLINSMVTNGLGLTTASGTD 871  
 DB 383 SPKALSPAPP-VDPSPCKGQASASRLSVSAAPFQSPGLSPSTYVGLGHTATATG 441  
 QY 872 ISSNSLKDCLPKSAOULKSFVKYGVGMATQLTSGAVWVQPNDSOLVQAGVSSISYSTP 931  
 DB 442 VSSS-----LPSAQLKSVFKYGVGMATQLTSGAVWVQPNDSOLVQAGVSSISYSTP 496  
 QY 932 NGQTRRYGENEKLPTYIKQKLCQLSSILLMFSNPPTNF 969  
 DB 497 DGQTRRYGENEKLPEYIKQKLCQLSSILLMFSNPPTNF 534  
 RESULT 7  
 Q9CVR6 PRELIMINARY; PRT; 324 AA.  
 AC Q9CVR6;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Serine/threonine kinase 18 (Fragment).  
 GN STK18.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi U., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,  
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kanakawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Caaveant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamliya W., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombauts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyszew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.",  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK006827; BAB24759.1; -.  
 DR HSSP; Q00534; 1B18.  
 DR MGD; MGI:101783; Stk18.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR002290; Ser Thr kinase.  
 DR InterPro; IPR001245; Tyr kinase.  
 DR InterPro; IPR002666; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; Kinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 FT ATP-binding; Transferase.  
 FT NON TER 324  
 SQ SEQUENCE 324 AA; 36568 MW; 04247065DFB198CC CRC64;  
 Query Match 31.2%; Score 1583.5; DB 11; Length 324;  
 Best Local Similarity 92.3%; Pred. No. 1.4e-95;  
 Matches 300; Conservative 15; Mismatches 9; Indels 1; Gaps 1;  
 QY 1 MATCIGERIDFKVGNLKGSPAGVYRAESITGTEVAIKMIDKAMYAKAGVQVONE 60  
 DB 1 MAACIGERIDFKVGNLKGSPAGVYRAESITGTEVAIKMIDKAMYAKAGVQVONE 60  
 QY 61 VKIHQCLKHSIIELVNYPEDSNVYLVLMKCHNGMNRVLRKRVYPSSENRARHFMQI 120  
 DB 61 VKIHQCLKHSIIELVNYPEDSNVYLVLMKCHNGMNRVLRKRVYPSSENRARHFMQI 120  
 QY 121 ITGMVLYSHSGIHRDLTSLNLTETNNMIKTADPGLAQOLKMPHKHNTLCTGTPYISP 180  
 DB 121 ITGMVLYSHSGIHRDLTSLNLTETNNMIKTADPGLAQOLKMPHKHNTLCTGTPYISP 180  
 QY 181 EIAITSAHGLSEDSVMSLGCMFYTLIIGRPPTDTYKNTLNKVLAADYEMPFLSTEARD 240  
 DB 181 EIAITSAHGLSEDSVMSLGCMFYTLIIGRPPTDTYKNTLNKVLAADYEMPFLSTEARD 240  
 QY 241 LIHQILRRNPADRLSLSSVLDHPFMSRNSYTSKDLGTWEDSIDSGHATITSAITASST 300  
 DB 241 LIHQILRRNPADRLSLSSVLDHPFMSRNSYTSKDLGTWEDSIDSGHATITSAITASST 300

QY 301 SIGSLFDRRLIGOLPKMTVP 325  
 DB 301 SIGSLLD-RRLVGOPLPKMTVP 324

## RESULT 8

Q7VVS3 PRELIMINARY; PRT; 711 AA.

AC 07VVS3  
 DT 01-JUN-2003 (TEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE Similar to serine/threonine kinase 18 (Fragment).  
 OS Brachydanio rerio (zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Body;  
 RA Strauberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC045434; AAH45434.1; -  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0007049; P:cell cycle; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000959; POLO box.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR PROSITE; PS50078; POLO BOX; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR KINASE.  
 KW NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 711 AA; 77340 MW; B8326280CCE72EAC CRC64;

Query Match 27 %; Score 1408.5; DB 13; Length 711;

Best Local Similarity 43.3%; Pred. No. 1.3e-83; Matches 334; Conservative 113; Mismatches 224; Indels 101; Gaps 24;

QY 230 MFSFLIEAKDLIQLRLRRPADRLSLVLDHPFMRNSSTKSDLGTYE-DSIDSGHA 288  
 DB 1 MFMHISAEKODLIQLRLRRPADRLSLVLDHPFMRNSSTKSDLGTYE-DSIDSGHA 60  
 QY 289 TISTATTAASSSTISGSLFDKRLIIGOLPKMTVPK--NKSSTDSSGSGDNGSFTT 345  
 DB 61 TISTATTAASSSTISGSLFDKRLIIGOLPKMTVPK--NKSSTDSSGSGDNGSFTT 113  
 QY 346 QW---GNQSTNSNGRGRVQDAEE-RPHRYLARAYSDRS--GTSNSGQAQTYTMC 399  
 DB 114 -QPNQSDQLSRWGVRIPIVAGDSGRPHSYLRRASSSDSAVGYSHNPQEA--LERC 169  
 QY 400 HSAEMLSVSKR-----SGGGENEERS-----PTDNANINFNFKETKSSSGSFER 446  
 DB 170 HSEEMLSGAGRLEPQSGYRNAPHGYSKIDRLSPVVKQAPNASSF--STHSTRQOM 227  
 QY 447 PNNQALSNHLCPGKTPFPADPTPQETVQVQFGLQINHLKRTTEYDISPN-RDPQ 505  
 DB 228 PQSQ-----TQWFSN-----DGVPKPADMS 249  
 QY 506 GHPLDQKQSKNMTDTXVKKNSDASDNHNSVQGMKMTALHKKPELIOECYFGSD 565  
 DB 250 GHSSSSGFSHSEKPIQOT-SCSDKPSGLHS--QQQPILEQ--HNPPGCRCD-AFVSG 302  
 QY 566 PLSEQSKTRGME---PPM--GYQNR-----TLRSTSPVLAHLRPIROKTKKAVASTLI 614  
 DB 303 HNSEPAYSADAGPCPCPLSKGKANTKKKQVCKKSPPLCAARLPIQKTKNAVASTLI 362  
 QY 615 DSEEVCELVKEVYASGEYKEVLQISSDQNTITTYVNGRGFPPLADRPSPPTDINSRY 674  
 DB 363 GNGEYCEMLKKGQAGQERVKEVLRISCDGSMVTYVQPNNEKGKFPVLDHPSPEDILIGS 422

QY 675 FDNLPKRYRKYOYASRFVOLYRSKPKITYFTFRYAKCILMENSPOADFEWFEYGVKIH 734  
 DB 423 FDLPEKRYRKYOYAKRFQOLYRSKPKITYFTFRYAKCILMENSPPDLEVCYDDAKTH 482  
 QY 735 KTEDFIQVIEKTKSYTLSESEVNSLKEIKMYMDHANEGRICLALR-SITSEERKT 793  
 DB 483 KTSQVRYVKEKSGSYTKGVKDVGLSGINPECRILYIELSEGHMCLSLAIAITAEQRSA 542  
 QY 794 RSPAPPIIIGRPGSTSSPKALSPPPSVDSNPTPTDRASFRNMWMSAASPQAPILNP 853  
 DB 543 KNTPEPPIITGRPVNPVPP---APAPSSSSC---RPAALAVHVLCSPPQHPQITP 595  
 QY 854 SWNTNGLGLITTAAGTSDISSNSLKQCLPKSAQLKSVFKVNGMAITOLTSQAVVQFND 913  
 DB 596 SMISYAGSDLTTSVAKGSSPVHKDERTYNSGVKLSITVPNIGWVSQLTGQVWQFND 655  
 QY 914 GSQLVQVGVSSISYTPNGQOTTRYGENEKLPDIYIKQKLCQLSSIIIMFSNP 965  
 DB 656 GSQLVQVGVSSISYTPNGQOTTRYGENEKLPDIYIKQKLCQLSSIIIMFSNP 707

## RESULT 9

Q97143 PRELIMINARY; PRT; 769 AA.

AC 097143  
 DT 01-MAY-1999 (TEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE SAK protein.  
 GN SAK OR CG7186.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkely;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Milos G.L.G.,  
 RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.B., Gary N.S., Galburt W.M., Glasner K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Herman T.J., Hernandez J.R., Houck J.,  
 RA Hotalik M., Houston K.A., Howland T.J., Wei M.-H., Ibeagwu K.A.,  
 RA Jaiswal M., Kalush F., Kapen G.H., Ke Z., Kemison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Modyarty C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Paclet J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J.D., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Hudson J.W., Dennis J.W.,  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF003594; AAF51737.1; -  
 DR EMBL; AF106952; AAD19607.1; -  
 DR HSSP; P00518; 1PHK  
 DR FlyBase; FBgn0026371; SAK.  
 DR GO; GO:000524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:Protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; F:Protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0007049; F:cell cycle; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000959; Polo\_box.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase.  
 DR Pfam; PF00069; pkinase\_1.  
 DR PRINTS; PR00109; TRYKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS50078; POLO\_BOX; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR KEGG; K00001; Kinase; Transferase.  
 SO SEQUENCE 769 AA; 8586 MW; F050BF60A5D94AA4 CRC64;

Query Match 26.4%; Score 1341.5; DB 5; length 769;  
 Best Local Similarity 33.0%; Pred. No. 3.6e-79;  
 Matches 321; Conservative 166; Mismatches 234; Indels 261; Gaps 24;

QY 6 GKEIEPKFKNGLKSGSPAGVTRAESITGLEYAIMIDKKMYAGVQVQNEVKIHC 65  
 DB 8 GETIEEYVQHLKSGKGFATVVKARCIHTHODVAIMDKLIGTGLTNRVQGEIHS 67  
 QY 66 QLKHPSELILYNYFEDNSNVYLVLEMGCHGEMNRVYKKNVKEPSENEAHFHQITGL 125  
 DB 68 RLKHPSELVLIYFPQANVYLVLELANHGLHRRYNNHARFTEASILIKQVAVGL 127  
 QY 126 YLHSHGILHDLITLSNLLITRNMTKIADFGLATQIKMPEHKGITLCTPNYISPEIATR 185  
 DB 128 YLHSHIMHARDISLISNLLISREHVKIADFGLATQIKRDERHMTMCGTPNYSISPEVSR 187  
 QY 186 SAHGESDVMSLGCMEFYTLTIGRPEDDTVTKANKVYLADYENPSTLSIAKDLIHQL 245  
 DB 188 TSHGLPADWAGVCMYTLTIGRPPEPTDAVOSTINKVMSRYIMPAHLSYEAQDLINKL 247  
 QY 246 LARNPADRLSSVLDHPEMSRNSSTKSKDLCTVEDSIDSGHATISTAITASSSTISGS 305  
 DB 248 LKKLPHERITLAVLCHEPMLKCS-----NGHSA----- 277  
 QY 306 LFDKRLILGQLPYNKMTYFPKNSSTD-----FSSSGDGNSTYTWGNGQETNSGRGV 360  
 DB 278 -----PALNWFQSQMSGDSGITTFASDSRSNQI-----RSEVNSGGPQGV 320  
 QY 361 IODAEERPHSRYLRAYSDDSGTNSQSAKTYMBCSHAEMLSVSKRGSGENERY 420  
 DB 321 LPOIREERKQVHKLPIY----- 337  
 QY 421 SPTDNANITFNFPEKETSSTSSGSEFRPPDNOALSNHLCPGKTPFPADPTPOTETVOQWF 480  
 DB 338 -----EQTGLF-----GQASGLAEP-----NWP 356  
 QY 481 GNLOINAHILKRTTEYDTSISPRNDPGHDLQKOTSXKAAWTOTKYKQSDASDNAAHSVQOQ 540  
 DB 357 GAASASATCMEA----- 368

QY 541 NTMKWTALHAKREIIQDECFVGSDDLSEOSKTRGMEPPWQYQNRNLSITPLVAHRLK 600  
 DB 369 -----GNVPSKQASL-----KEDRISVPLATYKILL 395  
 QY 601 PIRQKTKAVVSLIDSEECVAVK--EYASQERYKVLQISSDGNITTYYPNGRGPP 658  
 DB 396 PTRYKTKAVIWSILNNGEVLAEFLKRPPTYNDRINDICRISDDQRIIYOPDPGRGLP 455  
 QY 659 LADRPSS---PTDNISRSYFNDLPEKYRWKYQYASRFQVLRKSPKITYFRYAKCIIM 715  
 DB 456 VREGPPDLQIPSGDCV--VNYDPLPSKMKKITYGARFGLVKSCKPVTYITSLKCOLM 514  
 QY 716 EUSPGADREWYDGVKTHKT--EDFIQVIEKTKSYTLKSSEVNSLKEEIKMYDHANE 774  
 DB 515 EFM--TDEIFRYSAKLTKTPSEGLKYDRGMILSDYSCSESRL-----IEHNE 565  
 QY 775 GHRICALESIIEBERKTRASAPPEIIGRRPGSTSPKALSPPSVDSNYPRDRASF 834  
 DB 566 CETHCVNISMALEFVACQDNQSC--FVYTIIGRPITD-----VQPAQRIDG--LRDTTN- 614  
 QY 835 NEMVMSAASPTQADILNPSWVNEGLITTTASGDTISSNSLKDCLPKSAQ--LKSVF 892  
 DB 615 ---IAFSTPKSNQSGI--NFSIST-----ISSTRMTSDPGTNCSSNMALAHQNIPIKRI 665  
 QY 893 VKNVGMATQVLSGAVWQVQNGSOLV---QAGVSSISTSPNGQTTRYGENEKLDPYI 948  
 DB 666 VPEIGATLHSHGVQVQVQVGVSVVPSMVGQ--GITTYQPNGTSTHFEGKDDLFPV 723  
 QY 949 KQKLQCLSSILL 960  
 DB 724 RDRVQQLINQL 735

RESULT 10  
 Q86NL8 PRELIMINARY; PRT; 769 AA.  
 AC Q86NL8;  
 DT 01-JUN-2003 (Tremblrel. 24, Created)  
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE R57013ep.  
 GN Drosophila melanogaster (Fruit fly).  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Y;  
 RA stapleton M., Broketein P., Hong L., Agbayan A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Drenek D., Fafan D., Friese E.,  
 RA George R., Gonzalez W., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Munro J., Paclet J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Ceoliker S.;  
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BT004846; AAC45202.1; -  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:Protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; F:Protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0007049; P:cell cycle; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000959; Polo\_box.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR InterPro; IPR008266; Tyr\_kinase\_AS.  
 DR Pfam; PF00069; pkinase\_1.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TRYKc; 1.





Db 118 YTSDDGQTRRYGENEKLEPIYIKOKLOCLISIIIMFSNPTNPF 159

RESULT 12

086HN7 PRELIMINARY; PRT; 978 AA.

AC 086HN7, 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Similar to Aeterea pectinifera (Starfish). Polo-like kinase.

OS Dictyostelium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.

NCBI\_Taxid=44689;

RP SEQUENCE FROM N.A.

RC STRAIN=AX4;

RX MEDLINE=22092622; PubMed=12097910;

RA Gloeocher G., Bichinger L., Szafranski K., Pachebat J., Dear P.,

RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,

RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;

RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";

RL Nature 418:79-85(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=AX4;

RA Baumgart C.;

RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC116957; AA052535.1; -

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.

DR GO; GO:0007049; P:cell cycle; IEA.

DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.

DR InterPro; IPR000719; ProC kinase.

DR InterPro; IPR002290; Ser Thr\_pkinase.

DR InterPro; IPR008271; Ser Thr\_pkinase.

DR InterPro; IPR001245; Tyr\_pkinase.

DR Pfam; PF00069; Pkinase; 1.

DR ProDom; PD000001; Prot\_kinase; 1.

DR SMART; SM00220; S\_TKc; 1.

DR SMART; SM00219; TyrKc; 1.

DR PROSITE; PS00078; POLO BOX; 2.

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.

DR PROSITE; PS00113; PROTEIN KINASE\_DOM; 1.

DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.

KW Kinase.

SQ SEQUENCE 978 AA; 113599 MW; B03A677136543152 CRC64;

Query Match 11.7%; Score 593.5; DB 5; Length 978;

Best Local Similarity 21.7%; Pred. No. 5e-30;

Matches 219; Conservative 166; Mismatches 303; Indels 323; Gaps 37;

QY 7 EKIEDFKVGNLIGKSFAGVRAESIHGCELEVAIKMIDPKAMVYKAGVQVONVEYKHCQ 66

DB 158 KILKEVROGEFLGKGFACVIMTEVENIRYIAAKIIRKSTLOKTRASKSEKIKHS 217

QY 67 LKHPBILLEYNFEESNYVYVLENGCHGENRMYKKNVKPSENEAHFMHQTITGMLY 126

DB 218 LSHENIVKFEHCFENEENYVYVLENGCHGENRMYKKNVKPSENEAHFMHQTITGMLY 276

QY 127 LHSHGILHRDLITLSMLLTNRNNIKIADRGATOLKMH-EKHYTLCTEYVYSIETATR 185

DB 277 LHNNNIHRDLKGLNPLID-NRKIKGLDGLST-KVHGGRKKTICCTPYVIAPELIDN 333

QY 186 S-ANGLESDDVSLGCMFTYLLIGRPPPTDVTYKNTLKNVLADEMP5--PLSIEAKDLI 242

DB 334 SNGHSYEVADVMSIGIILYTLIGKPPFESDVKHYGRIKONQSFPEPIIISHGKSLI 393

QY 243 HQILRRNADLSSVLDHPFMSNNSSTKSDIGTVEDSIDSGHANTSTAISSSTSI 302

DB 394 ISILNPVEORPNLTQILIEHDFYSPIPK-----YLPVSSILVTAPSQSTI 439

QY 303 SGLFDRKRLILGQPLNMTVFPKNSSTDESSGDSNFPYONGNETSNGRGLVQ 362

DB 440 NQNM-----GRPLSEKTNIV-----NQQLQAGTSSPTKNNHHYQ 476

QY 363 DAERPHSYLRRAVSSDSGTSNSQSAQTYMERCHSAELSYSKSGGSENERISP 422

DB 477 QIQOQPOQOY-----NNNYQSFSPKKQINNNNNNNNNNNNNNNNNNNNN 523

QY 423 TNNNANIFNFEKTSSSSGSFEPDPNQLSNHLCPGKTPPADPTPOTETVOQWFGN 482

DB 524 NNNNLLQOYVY-----SNNNINNNNNNNNNNNNNNNNNNNNNNNNNNNNN 564

QY 483 LQ-INAHILKTEYDISIPNDPQGHPLQKDTSKNMTDTKYNK-SDASDNHVSQOQ 540

DB 565 VENDDPHYRKLRLKEMEN-----DLK-----TOLLIKOQYTNMENNQQOQOQ 609

QY 541 NTMKYMTALHSPREIIQOECVFGSDPLSEQSKTRGMEPPWGYQNRTLASITSPVAHRLK 600

DB 610 -----QQO-----

QY 601 PIRQTKKAVSILDSSEVCELVKEVASQBYKVELQISSDGNITTIYPNGRGFP-- 658

DB 613 -----QQOQQOQQOQRVNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 648

QY 659 -----LADRPSPPTDNISRYSPNLEPKYTRKQYASRFVQVRSKSPKITTFTR 708

DB 649 ELETYIANNNHISDSFPVSSNN-----NYP-----QQIQOQOQ----- 680

QY 709 YAKCIAMENSPGADFEVWFYDQ-----VKIKHTEDFIQLEKTKSYTLK----- 753

DB 681 -----NPNNEFYLGMPNNLVYISQVADF--TNKYLGLAVLSYGVAFND 724

QY 754 SESEVNSLKEIKYIMMDHA-----NEGRICLALSIISEERKTRSAAPF----- 799

DB 725 STKIYVLIIESIAYMEHAKGTGDGRRVLANVCOQHDPDQKVTLLIKYFLNHTNSDTT 784

QY 800 PIIIGRKPGSTSPKALSPPSVDSNYPTDRBASFNRMWMSAASPIQAILNSWYTN 859

DB 785 NLLI--NTGATSS-----SINN-----NNNNNN-- 807

QY 860 GLGLITTSAGTSDISSNKLCKLPKSAQLKSVFK-----NVGAWQTLGSAVWQFND 914

DB 808 -----VTNN 853

QY 915 SOLVVOAGVSSISYTSNGO-----TTRYGEN--EKLPDYIKOKLOCLISIL 959

DB 854 SRIIV-SSKDMVTFVPRGQIITGTLNPFYNGDKIKSEKIKYIGTILSNML 903

RESULT 13

Q7ZX26 PRELIMINARY; PRT; 663 AA.

AC 07ZX26, 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Similar to serum-inducible kinase.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;

OC Xenopodidae; Xenopus.

NCBI\_Taxid=8355;

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RA Klein S., Strausberg R.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC045272; AA045272.1; -

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

DR GO; GO:0007049; P:cell cycle; IEA.

DR GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000959; POLO\_box.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR002230; Ser\_thr\_pkinase.  
 DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro: IPR01245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00659; POLO\_box; 2.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS50078; POLO\_BOX; 2.  
 DR PROSITE: PS50107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR KINASE.  
 KW KINASE.  
 SQ SEQUENCE 663 AA; 76076 MW; 2E6B9A37C37639C CRC64;  
 Query Match 11.7%; Score 592; DB 13; Length 663;  
 Best Local Similarity 33.1%; Pred. No. 3.5e-30;  
 Matches 119; Conservative 73; Mismatches 156; Indels 12; Gaps 4;  
 QY 15 GNILGKSPAGYVRASIHITGLEVAIKMIDKKAMYGAVQVQVNEVKIHCOLKPSILE 74  
 DB 66 GKVLLGKGFARCYEMTDLTNNKIYAKIIPHSKVSXPHQREKIDKEIELHRLHHRHVQ 125  
 QY 75 LNYFEDSNVYVLVLEMCNGEMNRYLKNRVKPSNEARHFMHOITGMLYLHSHGILH 134  
 DB 126 FHHFEDADNIVYIPLFLCSKSLAHWKAR-HTLLEPEVRYLRQLLSGLKYLHQGILH 184  
 QY 135 RDLTSLNLLTNNMIKIDFGLATOLKMPHKKHTLCTPNYISPELATRAGHLESNV 194  
 DB 185 RDLKGNFPIITNMELKVGDFGLAARLEPEQRKRTICGTPNYLSBEVLAKQHGESDI 244  
 QY 195 NSLIGCMFYTLIGRPDPDVTYNTLNKVLADYEMPSFLSIEADOLIHQLRRPADRL 254  
 DB 244 MVLGCMYTMVLGSRPFETTNLKETRYCRREARYSLPSSLMTSKGLIASMLSRNEDRP 304  
 QY 255 SLSSVLDHPFMSNSTSKDGLTVEDSIDSGHATISTAITASSSTIGSLPDKRLLI 314  
 DB 305 HLDLIMQHDFFTOGFTPELPLTCCTADPFLHSPAKNPFKKYAAALFGGKKEKSYLD 364  
 QY 315 GQPLPKMTVPFNK-----KSTDPSSSGDGNSTYQWQNGQETNSGGRGVIODAERPH 369  
 DB 365 NH---NKL---PKDEEITYKLRODLQKTSISHQLNPRDDEIKNISKSDVLMKADKH 418  
 RESULT 14  
 Q96CV1 PRELIMINARY; PRT; 646 AA.  
 AC 096CV1; 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN [1]\_SEQUENCE FROM N.A.  
 RP TISSUE=Brain;  
 RA Strauberg R.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: BC013899; AAH1389.1; -  
 DR GO:0005524; F:ATP binding; IEA.  
 DR GO:0006474; P:protein serine/threonine kinase activity; IEA.  
 DR GO:0016740; F:transferase activity; IEA.  
 DR GO:0007049; P:cell cycle; IEA.  
 DR GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000959; POLO\_box.  
 DR InterPro: IPR000719; Prot\_kinase.

DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00659; POLO\_box; 2.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR PROSITE: PS50078; POLO\_BOX; 2.  
 DR PROSITE: PS50107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR Hypothetical protein; ATP-binding; Kinase.  
 KW Serine/threonine-protein kinase; transferase.  
 SQ SEQUENCE 646 AA; 71628 MW; 3242AD0D900865BE CRC64;  
 Query Match 11.6%; Score 590.5; DB 4; Length 646;  
 Best Local Similarity 37.5%; Pred. No. 4.2e-30;  
 Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;  
 QY 15 GNILGKSPAGYVRASIHITGLEVAIKMIDKKAMYGAVQVQVNEVKIHCOLKPSILE 74  
 DB 65 GNILGKGFARCYEMTDLTNNKIYAKIIPHSKVSXPHQREKIDKEIELHRLHHRHVQ 124  
 QY 75 LNYFEDSNVYVLVLEMCNGEMNRYLKNRVKPSNEARHFMHOITGMLYLHSHGILH 134  
 DB 125 FHHFEDADNIVYIPLFLCSKSLAHWKAR-HTLLEPEVRYLRQLLSGLKYLHQGILH 183  
 QY 135 RDLTSLNLLTNNMIKIDFGLATOLKMPHKKHTLCTPNYISPELATRAGHLESNV 194  
 DB 184 RDLKGNFPIITNMELKVGDFGLAARLEPEQRKRTICGTPNYVAEVLRLQHGELAV 243  
 QY 195 NSLIGCMFYTLIGRPDPDVTYNTLNKVLADYEMPSFLSIEADOLIHQLRRPADRL 254  
 DB 244 MVLGCMYTMVLGSRPFETTNLKETRYCRREARYSLPSSLMTSKGLIASMLSRNEDRP 303  
 QY 255 SLSSVLDHPFMSNSTSKDGLTVEDSIDSGHATISTAITASSSTIGSLPDKRLLI 314  
 DB 304 SIDQLRHDFPK-----GYTPDRL-----DISCVTPDLPMPARSLFAKYK 349  
 QY 305 SLFGRK 311  
 DB 350 SLFGRK 356  
 RESULT 15  
 Q8K0J7 PRELIMINARY; PRT; 504 AA.  
 AC 08K0J7; 01-OCT-2002 (TREMBLrel. 22, Created)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_Taxid=10090;  
 RN [1]\_SEQUENCE FROM N.A.  
 RP TISSUE=Colon;  
 RA Strauberg R.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 CC EMBL: BC031180; AAH31180.1; -  
 DR MGD; MG1:109604; Chk; -  
 DR GO:0005524; F:ATP binding; IEA.  
 DR GO:0006474; P:protein serine/threonine kinase activity; IEA.  
 DR GO:0004713; P:protein-tyrosine kinase activity; IEA.  
 DR GO:0016740; F:transferase activity; IEA.  
 DR GO:0007049; P:cell cycle; IEA.  
 DR GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000959; POLO\_box.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR008271; Ser\_thr\_pkin\_AS.

DR InterPro; IPR001245; Tyr\_kinase.  
DR Pfam; PF00069; kinase; 1.  
DR Pfam; PF00659; POLO box; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; TYKc; 1.  
DR PROSITE; PS50078; POLO\_BOX; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KM Hypothetical protein: ATP-binding; Transferase.  
SQ SEQUENCE 504 AA; 55546 MW; FC4BD79BBD273DB9 CRC64;

Query Match 11.6%; Score 587.5; DB 11; Length 504;  
Best Local Similarity 36.8%; Pred. No. 4.5e-30;

Matches 113; Conservative 64; Mismatches 105; Indels 25; Gaps 4;

QY 15 GNLLGSGSPAGVYRAESIRHGLEVALKMDKKMYKAGVQVQNEVKIHQQLKPSILE 74  
DB 66 GRLLGKGGFARCYEATDTEGIAVAVKIPQSRVAKPHQREKILNEIELHRDLQHRHIVR 125  
QY 75 LNVYFEDSNVYVLVEMCHNGEMNRYLKQNVKPFSENEAHEFMHQLITGMLYHSHGILH 134  
DB 126 FSHHFFEDADNIVYFLELCGRKSLAHYWKAR-HTLLEPEVRYVLRQILSGKYLHQRGILH 184  
QY 135 RDIITLSNLLITRNMTIKIADFGIATQLKMPHEKHYTLCTPNYISPEIATRSNAGLESYV 194  
DB 185 RDLKLGNFPTIDMELKVGDFGLARLEPPEQGRKTTICGTPNVVAPEVLLRQGHPEADV 244  
QY 195 WSLGCMFYTLILGRPEPDITVKNLTKNVVLADYEMPSFLSTFAKDLIHQLRRNPADRL 254  
DB 245 WSLGCMFYTLILGSPPEPTADLKETRCIKQVHYTLPASLSLPARQLAAILRASPRDRP 304  
QY 255 SLSSVLDPHPMSRNSSTKSKDLGTVEDSIDSGHATISTAITASSST-----SIGG 304  
DB 305 SIEQILRHDFETK-----GYTPDRL-----PVSSCVTVPDLTTPNPASLPFAKVTK 350  
QY 305 SLPEDKRR 311  
DB 351 SLFGRRK 357

Search completed: September 23, 2004, 21:02:04  
Job time : 69 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 21:00:41 ; Search time 69 Seconds  
(without alignments)  
4520.463 Million cell updates/sec

Title: US-10-026-021-2

Perfect score: 5078

Sequence: 1 MATCGEKIEDPKVGNLTKGK.....KLQCLSSILMSPNTPNPH 970

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 32158718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5078	100.0	970	14	US-10-026-021-2
2	5074	99.9	970	16	US-10-408-765A-1916
3	4861	95.7	928	12	US-10-425-114-37528
4	1973	38.9	379	14	US-10-026-021-3
5	590.5	11.6	607	10	US-09-769-970-15
6	590.5	11.6	607	14	US-10-108-580-2
7	590.5	11.6	607	14	US-10-204-041-16
8	590.5	11.6	607	14	US-10-620-052A-28
9	590.5	11.6	607	14	US-10-755-889-4
10	589.5	11.6	373	16	US-10-026-021-4
11	580.5	11.4	253	16	US-10-620-052A-43
12	568.5	11.2	1462	16	US-10-618-581-15
13	560.5	11.0	400	14	US-10-026-021-5
14	560.5	11.0	685	9	US-09-771-161A-249
15	560.5	11.0	685	9	US-09-771-161A-250

16	560.5	11.0	685	9	US-09-771-161A-251	Sequence 251, App
17	560.5	11.0	685	10	US-09-769-970-1	Sequence 1, Appl
18	560.5	11.0	685	12	US-10-260-708-69	Sequence 69, Appl
19	560.5	11.0	685	14	US-10-024-298A-101	Sequence 101, App
20	560.5	11.0	685	14	US-10-042-211A-101	Sequence 101, App
21	560.5	11.0	685	16	US-10-617-217A-101	Sequence 101, App
22	560.5	11.0	753	15	US-10-264-049-3124	Sequence 3124, App
23	544.5	10.7	1246	15	US-10-369-493-6585	Sequence 6585, App
24	538.5	10.6	896	16	US-10-437-963-136508	Sequence 136508, App
25	538	10.6	603	9	US-09-771-161A-214	Sequence 214, App
26	536	10.6	603	16	US-10-188-832-110	Sequence 110, App
27	536	10.6	603	16	US-10-408-765A-2279	Sequence 2279, App
28	536	10.6	629	12	US-10-425-114-37525	Sequence 37525, A
29	535.5	10.5	1038	12	US-10-424-599-274878	Sequence 274878, A
30	533	10.5	603	12	US-10-406-901-2	Sequence 2, Appl
31	533	10.5	603	14	US-10-171-311-186	Sequence 186, App
32	530	10.4	367	14	US-10-026-021-6	Sequence 6, Appl
33	527	10.4	461	16	US-10-767-701-36696	Sequence 36696, A
34	526.5	10.4	1518	9	US-09-801-366-152	Sequence 152, App
35	526.5	10.4	1518	15	US-10-369-493-22243	Sequence 22243, A
36	526.5	10.4	1518	16	US-10-618-581-11	Sequence 11, Appl
37	525	10.3	353	16	US-10-664-421-124	Sequence 124, App
38	524	10.3	523	12	US-10-183-687-258	Sequence 258, App
39	520	10.2	493	16	US-10-437-963-184443	Sequence 184443, A
40	517.5	10.2	1349	16	US-10-618-581-16	Sequence 16, Appl
41	516.5	10.2	329	9	US-09-925-300-1268	Sequence 1268, App
42	516.5	10.2	832	9	US-09-919-585-21	Sequence 21, Appl
43	516.5	10.2	832	15	US-10-464-939-2	Sequence 2, Appl
44	514.5	10.1	683	15	US-10-369-493-2188	Sequence 2188, App
45	510.5	10.1	402	12	US-10-087-684-51	Sequence 51, Appl

## ALIGNMENTS

RESULT 1  
US-10-026-021-2  
; Sequence 2, Application US/10026021  
; Publication No. US20030027756A1  
; GENERAL INFORMATION:  
; APPLICANT: Hitoenhi, Yasumichi  
; APPLICANT: Demo, Susan  
; APPLICANT: Denkins, Yonchu  
; APPLICANT: Rigel Pharmaceuticals, Inc.  
; TITLE OF INVENTION: SAK: Modulation of Cellular Proliferation for  
; TITLE OF INVENTION: Treatment of Cancer  
; FILE REFERENCE: 021044-001210US  
; CURRENT APPLICATION NUMBER: US/10/026,021  
; CURRENT FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/309,632  
; PRIOR FILING DATE: 2001-08-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 970  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human SAK serine/threonine kinase  
; US-10-026-021-2

Query Match 100.0%; Score 5078; DB 14; Length 970;  
Best local similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 970; Conservative 0;

QY 1 MATCGEKIEDPKVGNLTKGSPAGYRAESHTGTEVAIKMIDKKAMTKAGVORVONE 60  
DB 1 MATCGEKIEDPKVGNLTKGSPAGYRAESHTGTEVAIKMIDKKAMTKAGVORVONE 60  
QY 61 VAIHCOLKRPSTLELYNYFEDSNYYVLVEHCHGEMRRYLKRRYKPFSENAHRHMOI 120  
DB 61 VAIHCOLKRPSTLELYNYFEDSNYYVLVEHCHGEMRRYLKRRYKPFSENAHRHMOI 120

```

QY 121 ITGMLYLHSHGILHARDLTLNSLLLTRNNMIKIADFGIATOLKMPHEKHYTLGTPNYISP 180
Db 121 ITGMLYLHSHGILHARDLTLNSLLLTRNNMIKIADFGIATOLKMPHEKHYTLGTPNYISP 180
QY 181 EIAIRSAHGLESDVWSLGCMEFYTLILGRPPDTDTVKNTLNKVVLADYEMPSFLISBEAKD 240
Db 181 EIAIRSAHGLESDVWSLGCMEFYTLILGRPPDTDTVKNTLNKVVLADYEMPSFLISBEAKD 240
QY 241 LHHQLLRNPADRLSLSSVLDPHPMSRNSSTKSGLGTVEDSIDSGHATITATITASSST 300
Db 241 LHHQLLRNPADRLSLSSVLDPHPMSRNSSTKSGLGTVEDSIDSGHATITATITASSST 300
QY 301 SISGSLFDRKRLILIGPLPNKMTVFPKXKSTDPSSGDSNGSFYTOGNOETSNISGRGV 360
Db 301 SISGSLFDRKRLILIGPLPNKMTVFPKXKSTDPSSGDSNGSFYTOGNOETSNISGRGV 360
QY 361 IQDAERPHSRYLRRAYSSDRSGTNSGQAKTYTMRCHSAEMLSVSRSGGGENEERY 420
Db 361 IQDAERPHSRYLRRAYSSDRSGTNSGQAKTYTMRCHSAEMLSVSRSGGGENEERY 420
QY 421 SPTDNNANIFNPFKEKTSSSGSGFERPDNNQALSNHLCPGKTPFPADPTPOTETVQOMF 480
Db 421 SPTDNNANIFNPFKEKTSSSGSGFERPDNNQALSNHLCPGKTPFPADPTPOTETVQOMF 480
QY 481 GNLQINAHILRKTTEYDVISPNRDPQGHDLQKDTSKNMTDTKVKKNSDASDNASHVKQ 540
Db 481 GNLQINAHILRKTTEYDVISPNRDPQGHDLQKDTSKNMTDTKVKKNSDASDNASHVKQ 540
QY 541 NTMKWTALHSHKPEIIQOECVFGSDPLSEOSKTRGMEPPWGYQNRILRSITSPVAHRLK 600
Db 541 NTMKWTALHSHKPEIIQOECVFGSDPLSEOSKTRGMEPPWGYQNRILRSITSPVAHRLK 600
QY 601 PIROKTKKAVSILDSEEVCELVKEYASQEVYKVLQISSDGNITITYYPNRGSGPPLA 660
Db 601 PIROKTKKAVSILDSEEVCELVKEYASQEVYKVLQISSDGNITITYYPNRGSGPPLA 660
QY 661 DRPSPDTNISRYSPDNLPKTYWKRYQYASRFVQLVRSKPKITYFTRVAKCILMENSFG 720
Db 661 DRPSPDTNISRYSPDNLPKTYWKRYQYASRFVQLVRSKPKITYFTRVAKCILMENSFG 720
QY 721 ADPEWVFYDGVKIKHTEDFIOVIEKTKGSKYTLKSSEVNSLKEEIKMYMDHANEGRICL 780
Db 721 ADPEWVFYDGVKIKHTEDFIOVIEKTKGSKYTLKSSEVNSLKEEIKMYMDHANEGRICL 780
QY 781 ALESIISEBERKTRSAFPFPIIIGRKPGSTSSPKALSPPSVDSNYPTRDRASFNRVMWH 840
Db 781 ALESIISEBERKTRSAFPFPIIIGRKPGSTSSPKALSPPSVDSNYPTRDRASFNRVMWH 840
QY 841 SAASPTQAPILNPSVNTNBSGLITTTASGTDISSNSLKDCLEKSAQLKSVFVKNVGNAT 900
Db 841 SAASPTQAPILNPSVNTNBSGLITTTASGTDISSNSLKDCLEKSAQLKSVFVKNVGNAT 900
QY 901 QUTSAGVWVQFNDGSQLVVOAGVSSISYTSPNGQTRRGENEKLDYIKOKIQCCSSTIL 960
Db 901 QUTSAGVWVQFNDGSQLVVOAGVSSISYTSPNGQTRRGENEKLDYIKOKIQCCSSTIL 960
QY 961 MFSNPTPNFH 970
Db 961 MFSNPTPNFH 970

```

## RESULT 2

```

US-10-408-765A-1916
; Sequence 1916, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Rahy, Bojin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.

```

```

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1916
; LENGTH: 970
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1916

```

Query Match 99.9%; Score 5074; DB 16; Length 970;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 969; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MATCIGEKIEDPKVGNLKGSGFAGVYRAESIHTGLAEVAIKMIDKKAMYKAGMYQVQNE 60
Db 1 MATCIGEKIEDPKVGNLKGSGFAGVYRAESIHTGLAEVAIKMIDKKAMYKAGMYQVQNE 60
QY 61 VKIHQLKHPSTILELYNPFEDSNYYLYVLEMGCHGENRKYLNRYKPSSEBARHFHQI 120
Db 61 VKIHQLKHPSTILELYNPFEDSNYYLYVLEMGCHGENRKYLNRYKPSSEBARHFHQI 120
QY 121 ITGMLYLHSHGILHARDLTLNSLLLTRNNMIKIADFGIATOLKMPHEKHYTLGTPNYISP 180
Db 121 ITGMLYLHSHGILHARDLTLNSLLLTRNNMIKIADFGIATOLKMPHEKHYTLGTPNYISP 180
QY 181 EIAIRSAHGLESDVWSLGCMEFYTLILGRPPDTDTVKNTLNKVVLADYEMPSFLISBEAKD 240
Db 181 EIAIRSAHGLESDVWSLGCMEFYTLILGRPPDTDTVKNTLNKVVLADYEMPSFLISBEAKD 240
QY 241 LHHQLLRNPADRLSLSSVLDPHPMSRNSSTKSGLGTVEDSIDSGHATITATITASSST 300
Db 241 LHHQLLRNPADRLSLSSVLDPHPMSRNSSTKSGLGTVEDSIDSGHATITATITASSST 300
QY 301 SISGSLFDRKRLILIGPLPNKMTVFPKXKSTDPSSGDSNGSFYTOGNOETSNISGRGV 360
Db 301 SISGSLFDRKRLILIGPLPNKMTVFPKXKSTDPSSGDSNGSFYTOGNOETSNISGRGV 360
QY 361 IQDAERPHSRYLRRAYSSDRSGTNSGQAKTYTMRCHSAEMLSVSRSGGGENEERY 420
Db 361 IQDAERPHSRYLRRAYSSDRSGTNSGQAKTYTMRCHSAEMLSVSRSGGGENEERY 420
QY 421 SPTDNNANIFNPFKEKTSSSGSGFERPDNNQALSNHLCPGKTPFPADPTPOTETVQOMF 480
Db 421 SPTDNNANIFNPFKEKTSSSGSGFERPDNNQALSNHLCPGKTPFPADPTPOTETVQOMF 480
QY 481 GNLQINAHILRKTTEYDVISPNRDPQGHDLQKDTSKNMTDTKVKKNSDASDNASHVKQ 540
Db 481 GNLQINAHILRKTTEYDVISPNRDPQGHDLQKDTSKNMTDTKVKKNSDASDNASHVKQ 540
QY 541 NTMKWTALHSHKPEIIQOECVFGSDPLSEOSKTRGMEPPWGYQNRILRSITSPVAHRLK 600
Db 541 NTMKWTALHSHKPEIIQOECVFGSDPLSEOSKTRGMEPPWGYQNRILRSITSPVAHRLK 600
QY 601 PIROKTKKAVSILDSEEVCELVKEYASQEVYKVLQISSDGNITITYYPNRGSGPPLA 660
Db 601 PIROKTKKAVSILDSEEVCELVKEYASQEVYKVLQISSDGNITITYYPNRGSGPPLA 660
QY 661 DRPSPDTNISRYSPDNLPKTYWKRYQYASRFVQLVRSKPKITYFTRVAKCILMENSFG 720
Db 661 DRPSPDTNISRYSPDNLPKTYWKRYQYASRFVQLVRSKPKITYFTRVAKCILMENSFG 720
QY 721 ADPEWVFYDGVKIKHTEDFIOVIEKTKGSKYTLKSSEVNSLKEEIKMYMDHANEGRICL 780
Db 721 ADPEWVFYDGVKIKHTEDFIOVIEKTKGSKYTLKSSEVNSLKEEIKMYMDHANEGRICL 780
QY 781 ALESIISEBERKTRSAFPFPIIIGRKPGSTSSPKALSPPSVDSNYPTRDRASFNRVMWH 840
Db 781 ALESIISEBERKTRSAFPFPIIIGRKPGSTSSPKALSPPSVDSNYPTRDRASFNRVMWH 840

```

```

QY 841 SAASPTQAPILINPMVTNEGILGTTTASGTDISSNLSKDCLPKSAQLLKSFVKNVGMAT 900
DB 841 SAASPTQAPILINPMVTNEGILGTTTASGTDISSNLSKDCLPKSAQLLKSFVKNVGMAT 900
QY 901 QLTSGAVWVQFNDGSQLVVOAGVSSISYTSPNGQTRRYGENEKLDPYIKQKQCLSSILL 960
DB 901 QLTSGAVWVQFNDGSQLVVOAGVSSISYTSPNGQTRRYGENEKLDPYIKQKQCLSSILL 960
QY 961 MFSNPTNPFH 970
DB 961 MFSNPTNPFH 970

```

## RESULT 3

```

US-10-425-114-37528
/ Sequence 37528; Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Zhou, Jindong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313) B
/ CURRENT APPLICATION NUMBER: US/10/425, 114
/ NUMBER OF SEQ ID NOS: 2003-04-28
/ SEQ ID NO 37528
/ LENGTH: 928
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB4119-112-84_FRI.pep
US-10-425-114-37528

```

```

Query Match      95.7%; Score 4861; DB 12; Length 928;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 927; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 43 IDKKAMVKKAGMVOVONEVYKHCQLKHPSTILELYNFBEDSNYYVLEMCNENRNYLK 102
DB 1 IDKKAMVKKAGMVOVONEVYKHCQLKHPSTILELYNFBEDSNYYVLEMCNENRNYLK 60
QY 103 NRVKPSSEBARHFMHQITGMVYLHSHGILHRDLTSLNLLTRNNNIKIADFGLATOLK 162
DB 61 NRVKPSSEBARHFMHQITGMVYLHSHGILHRDLTSLNLLTRNNNIKIADFGLATOLK 120
QY 163 MPHEKAYTLCGTPNYTISPETATRSANGLESVDWVSLGCMFTYLLIGRPPPTDTVVKATLTK 222
DB 121 MPHEKAYTLCGTPNYTISPETATRSANGLESVDWVSLGCMFTYLLIGRPPPTDTVVKATLTK 180
QY 223 VVLADYEMPSFSLIEKDLIHLRRNPDRLSLSVLDHPFMSRSTKSKDGLTVEDS 282
DB 181 VVLADYEMPSFSLIEKDLIHLRRNPDRLSLSVLDHPFMSRSTKSKDGLTVEDS 240
QY 283 IDSGHATITATYASSSTSISGSLPDKRRLLIGQPLPNKATVPPKXKSTDFSSSGDGNS 342
DB 241 IDSGHATITATYASSSTSISGSLPDKRRLLIGQPLPNKATVPPKXKSTDFSSSGDGNS 300
QY 343 FTYQMGNOTSNGSGRGVIODAEERPHSRYLRAVYSSDNGSGTNSGSOATYMERCHSA 402
DB 301 FTYQMGNOTSNGSGRGVIODAEERPHSRYLRAVYSSDNGSGTNSGSOATYMERCHSA 360
QY 403 EMLSVKSGSGGSENERYSPTDNNANIENFKEKETSSSSGSPRPDNNQQLSNHLCPGKT 462
DB 361 EMLSVKSGSGGSENERYSPTDNNANIENFKEKETSSSSGSPRPDNNQQLSNHLCPGKT 420
QY 463 PPPFADPTPTQTEFVQMGFGLQINAHLRKTTEDYSISPNDFOGHDLQKDTSKNMTDT 522
DB 421 PPPFADPTPTQTEFVQMGFGLQINAHLRKTTEDYSISPNDFOGHDLQKDTSKNMTDT 480

```

```

QY 523 KYKNSDASDNASHYKQNTMKMTALHSEKPEIIQOECVFGSDPLSEQSKTRGMEPPMGY 582
DB 481 KYKNSDASDNASHYKQNTMKMTALHSEKPEIIQOECVFGSDPLSEQSKTRGMEPPMGY 540
QY 583 QNRTLSITSPLVAHLKPIROKTKKAVVSIIDSEEVCELVKEYASQEVYKVLQISSD 642
DB 541 QNRTLSITSPLVAHLKPIROKTKKAVVSIIDSEEVCELVKEYASQEVYKVLQISSD 600
QY 643 GNTTITTYPNGRGKPLADRPSPPTDNIISRYSPDNIPEKYMRYQYASFPVQLVNSKSK 702
DB 601 GNTTITTYPNGRGKPLADRPSPPTDNIISRYSPDNIPEKYMRYQYASFPVQLVNSKSK 660
QY 703 ITYFTRACIIMENSPGADFEWPFYDVKIHKTEDFIVIEKTKGSYLLKSESEVNSLK 762
DB 661 ITYFTRACIIMENSPGADFEWPFYDVKIHKTEDFIVIEKTKGSYLLKSESEVNSLK 720
QY 763 BEIKMYMDHANGHRIICALLESIISEERKTSAPFPPIIIGRKPGSTSPKALSPPSV 822
DB 721 BEIKMYMDHANGHRIICALLESIISEERKTSAPFPPIIIGRKPGSTSPKALSPPSV 780
QY 823 DSNYPTRDASFRMYMHSNAASTQAPILINPMVTNEGILGTTTASGTDISSNLSKDCLP 882
DB 781 DSNYPTRDASFRMYMHSNAASTQAPILINPMVTNEGILGTTTASGTDISSNLSKDCLP 840
QY 883 KSAQLLKSFVKNVGMATQLTSGAVWVQFNDGSQLVVOAGVSSISYTSPNGQTRRYGENE 942
DB 841 KSAQLLKSFVKNVGMATQLTSGAVWVQFNDGSQLVVOAGVSSISYTSPNGQTRRYGENE 900
QY 943 KLPDYIKQKQCLSSILLMFSNPTNPFH 970
DB 901 KLPDYIKQKQCLSSILLMFSNPTNPFH 928

```

## RESULT 4

```

US-10-026-021-3
/ Sequence 3; Application US/10026021
/ Publication No. US20030027756A1
/ GENERAL INFORMATION:
/ APPLICANT: Hitoshi, Yasumichi
/ APPLICANT: Demo, Susan
/ APPLICANT: Jenkins, Yonchu
/ APPLICANT: Rigel Pharmaceuticals, Inc.
/ TITLE OF INVENTION: SAK: Modulation of Cellular Proliferation for
/ TITLE OF INVENTION: Treatment of Cancer
/ FILE REFERENCE: 021044-001210US
/ CURRENT APPLICATION NUMBER: US/10/026, 021
/ PRIOR FILING DATE: 2002-06-25
/ PRIOR APPLICATION NUMBER: US 60/309, 632
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 3
/ LENGTH: 379
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: DOMAIN
/ LOCATION: (1)..(379)
/ OTHER INFORMATION: SAK serine/threonine kinase domain
US-10-026-021-3

```

```

Query Match      38.9%; Score 1973; DB 14; Length 379;
Best Local Similarity 100.0%; Pred. No. 3.7e-127;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MATCIGKEIDEPFVGNLKGSGFAGYVRAESIHTGLEVAIKMIDKKAMTKAGVQVQNE 60
DB 1 MATCIGKEIDEPFVGNLKGSGFAGYVRAESIHTGLEVAIKMIDKKAMTKAGVQVQNE 60
QY 61 VKIHQCLKHPSIILELYNFBEDSNYYVLEMCNENRNYLKNRVKPSSEBARHFMHQI 120
DB 61 VKIHQCLKHPSIILELYNFBEDSNYYVLEMCNENRNYLKNRVKPSSEBARHFMHQI 120

```



QY 121 INGMYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHYTLCTGPNYISP 180  
 DB 121 INGMYLHSHGILHRDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHYTLCTGPNYISP 180  
 QY 181 EIATISAHGLESVDWSLGCMFYTLIGRPPTDTYKNTLANKVVLADYEMPSFLISIAKD 240  
 DB 181 EIATISAHGLESVDWSLGCMFYTLIGRPPTDTYKNTLANKVVLADYEMPSFLISIAKD 240  
 QY 241 LHHLLRRPADRLSLSSVLDHPFMSRNSSTSKDGLTVEDSIDSGHATISATITASSST 300  
 DB 241 LHHLLRRPADRLSLSSVLDHPFMSRNSSTSKDGLTVEDSIDSGHATISATITASSST 300  
 QY 301 SISGLFDRKRLLIQPLPNKMTVPFPKNSSTDPSSSGDSGNSFYTQWNGOETSNSGRGRV 360  
 DB 301 SISGLFDRKRLLIQPLPNKMTVPFPKNSSTDPSSSGDSGNSFYTQWNGOETSNSGRGRV 360  
 QY 361 IODAEERPHSRYLRRAYSS 379  
 DB 361 IODAEERPHSRYLRRAYSS 379

## RESULT 5

US-09-769-970-15  
 ; Sequence 15, Application US/09769970  
 ; Publication No. US20030170219A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bandman, Olga  
 ; Hillman, Jennifer L.  
 ; Corley, Neil C.  
 ; Guegler, Karl G.  
 ; Lal, Preeti  
 ; Goli, Surya K.  
 ; Shah, Purvi  
 ; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
 ; KINASES  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/769, 970  
 ; FILING DATE: 24-Jan-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 09/272,796  
 ; FILING DATE: <Unknown>  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0321 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; TELEX: <Unknown>  
 ; INFORMATION FOR SEQ ID NO: 15:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 607 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GenBank  
 ; CLONE: 1827450

SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
 ; US-09-769-970-15

Query Match 11.6%; Score 590.5; DB 10; Length 607;  
 Best Local Similarity 37.5%; Pred. No. 7,1e-32;  
 Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;

QY 15 GNLLKSGSPAGVYRAESIHTGLEVAIKMIDKKAMTKAGMVRQVNEVKIHQOLKHPSTLE 74  
 DB 26 GRLLKSGGFARCYETATDTEGSAVAVKVIPOSRAVAKPQREKINIEIHLRDLQHRHIVR 85  
 QY 75 LNYVEDSNVYVIVLEWCHNGEMNRVYLKNRVKPFSENEARHPMQIITGMLYLHSHGILH 134  
 DB 86 FSHHEDNDNITYIFELCSRSKSLAHWIKAR-HTLLEPVRVYLRQILSLGKLTHORGLIH 144  
 QY 135 RDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHYTLCTGPNYISPELATISAHGLESVDV 194  
 DB 145 RDLKLGNFPIENNELKVGDFGLAARLEPPQRKKTIGTNTVYAPVLLRQGHGPEADV 204  
 QY 195 WSLGCMFYTLIGRPPTDTYKNTLANKVVLADYEMPSFLISIAKDILHOLLRRPADRL 254  
 DB 205 WSLGCMVYTLICGSPFPETADLKETYRCIKQVHYTLPLASLSIPARQLAAILRASPDRP 264  
 QY 255 SISGLFDRKRLLIQPLPNKMTVPFPKNSSTDPSSSGDSGNSFYTQWNGOETSNSGRGRV 304  
 DB 265 SIDQILRHDFFTK-----GYPPDRL-----PISCCVTPDULTPPNPARSLPAKYTK 310  
 QY 305 SLFDRKR 311  
 DB 311 SLFGRKK 317

## RESULT 6

US-10-108-580-2  
 ; Sequence 2, Application US/10108580  
 ; Publication No. US20030077681A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cogswell, John  
 ; TITLE OF INVENTION: PLK3 PROTEIN-PROTEIN INTERACTIONS  
 ; FILE REFERENCE: P04458  
 ; CURRENT APPLICATION NUMBER: US/10/108,580  
 ; CURRENT FILING DATE: 2002-03-28  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 607  
 ; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 ; US-10-108-580-2

Query Match 11.6%; Score 590.5; DB 14; Length 607;  
 Best Local Similarity 37.5%; Pred. No. 7,1e-32;  
 Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;

QY 15 GNLLKSGSPAGVYRAESIHTGLEVAIKMIDKKAMTKAGMVRQVNEVKIHQOLKHPSTLE 74  
 DB 26 GRLLKSGGFARCYETATDTEGSAVAVKVIPOSRAVAKPQREKINIEIHLRDLQHRHIVR 85  
 QY 75 LNYVEDSNVYVIVLEWCHNGEMNRVYLKNRVKPFSENEARHPMQIITGMLYLHSHGILH 134  
 DB 86 FSHHEDNDNITYIFELCSRSKSLAHWIKAR-HTLLEPVRVYLRQILSLGKLTHORGLIH 144  
 QY 135 RDLTSLNLLTRNNMIKIADFGLATOLKMPHEKHYTLCTGPNYISPELATISAHGLESVDV 194  
 DB 145 RDLKLGNFPIENNELKVGDFGLAARLEPPQRKKTIGTNTVYAPVLLRQGHGPEADV 204  
 QY 195 WSLGCMFYTLIGRPPTDTYKNTLANKVVLADYEMPSFLISIAKDILHOLLRRPADRL 254  
 DB 205 WSLGCMVYTLICGSPFPETADLKETYRCIKQVHYTLPLASLSIPARQLAAILRASPDRP 264  
 QY 255 SISGLFDRKRLLIQPLPNKMTVPFPKNSSTDPSSSGDSGNSFYTQWNGOETSNSGRGRV 304  
 DB 265 SIDQILRHDFFTK-----GYPPDRL-----PISCCVTPDULTPPNPARSLPAKYTK 310

QY 305 SLFDRK 311  
DB 311 SLFGRK 317

## RESULT 7

US-10-204-041-16  
Sequence 16, Application US/10204041  
Publication No. US2003017643A1  
GENERAL INFORMATION:  
APPLICANT: STEIN-GERLACH, MATTHIAS  
APPLICANT: SALASIDIS, KONSTANTINOS  
APPLICANT: BACHER, GERARD  
APPLICANT: MULLER, STERAN  
TITLE OF INVENTION: Pyridylpyrimidine Derivatives as Effective Compounds Against Prio  
FILE REFERENCE: AXM-007.1P US  
CURRENT APPLICATION NUMBER: US/10/204,041  
PRIOR FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: EP 0111858.5  
PRIOR FILING DATE: 2001-05-16  
PRIOR APPLICATION NUMBER: PCT/EP02/05420  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 16  
LENGTH: 607  
TYPE: PR  
ORGANISM: Homo sapiens  
US-10-204-041-16

Query Match 11.6%; Score 590.5; DB 14; Length 607;  
Best Local Similarity 37.5%; Pred. No. 7.1e-32;  
Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;

QY 15 GNLLGSGFAGVYRABSIHTGLEVAIKMIDKKAMYGAVORVONEVKIHCQKPSILE 74  
DB 26 GNLLGSGFAGVYRABSIHTGLEVAIKMIDKKAMYGAVORVONEVKIHCQKPSILE 74  
QY 75 LNYFEDSNVYVLVLEMGCHGEMNRYLKNRVKPFSENEARHFMHQIITGMLYHSHGILH 134  
DB 86 FSHHPEADNIIYIFELGSRKSLAHIMWKAR-HTLLEPVRYVLRQLSLGKYLHGQGIH 144  
QY 135 RDLTSLNLLITRNWNIKIADFGLATOLKMPHEKHYTLCTGPNYISPEIATRSAGHESDV 194  
DB 145 RDLKGNFETITENMEIKVDFGLAARLEBPQRKKTICGTPNVVAEBVLLRQGHGEADV 204  
QY 195 WSLGCMFYTLTGRPPFDIVTNTLKNVVLADYEMPSFLISAKDLIHQLRRNPADRL 254  
DB 205 WSLGCMFYTLTGRPPFDIVTNTLKNVVLADYEMPSFLISAKDLIHQLRRNPADRL 254  
QY 255 SLSSVLDHPFMSRNSSTKSLDGTVEDSIDSGHATISTATSSST-----SIG 304  
DB 265 SLDDQLRHDFPK-----GYTPDRL-----PSSCTVVDLTPNPABSLPAKVT 310  
QY 305 SLFDRK 311  
DB 311 SLFGRK 317

## RESULT 8

US-10-620-052A-28  
Sequence 28, Application US/10620052A  
Publication No. US20040126784A1  
GENERAL INFORMATION:  
APPLICANT: Hitoebl, Yasumichi  
APPLICANT: Jenkine, Yochiu  
APPLICANT: Markovtsov, Vadim  
APPLICANT: Rigel Pharmaceuticals, Inc.  
TITLE OF INVENTION: Modulators of Cellular Proliferation  
FILE REFERENCE: 021044-004010US  
CURRENT APPLICATION NUMBER: US/10/620,052A

CURRENT FILING DATE: 2003-07-14  
PRIOR APPLICATION NUMBER: US 60/395,443  
PRIOR FILING DATE: 2002-07-12  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 28  
LENGTH: 607  
TYPE: PR  
ORGANISM: Homo sapiens

OTHER INFORMATION: cytokine-inducible kinase (CKK) serine threonine  
OTHER INFORMATION: kinase, proliferation-related kinase (PRK),  
OTHER INFORMATION: polo-like kinase 3 (PLK3)  
US-10-620-052A-28

Query Match 11.6%; Score 590.5; DB 16; Length 607;  
Best Local Similarity 37.5%; Pred. No. 7.1e-32;  
Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;

QY 15 GNLLGSGFAGVYRABSIHTGLEVAIKMIDKKAMYGAVORVONEVKIHCQKPSILE 74  
DB 26 GNLLGSGFAGVYRABSIHTGLEVAIKMIDKKAMYGAVORVONEVKIHCQKPSILE 74  
QY 75 LNYFEDSNVYVLVLEMGCHGEMNRYLKNRVKPFSENEARHFMHQIITGMLYHSHGILH 134  
DB 86 FSHHPEADNIIYIFELGSRKSLAHIMWKAR-HTLLEPVRYVLRQLSLGKYLHGQGIH 144  
QY 135 RDLTSLNLLITRNWNIKIADFGLATOLKMPHEKHYTLCTGPNYISPEIATRSAGHESDV 194  
DB 145 RDLKGNFETITENMEIKVDFGLAARLEBPQRKKTICGTPNVVAEBVLLRQGHGEADV 204  
QY 195 WSLGCMFYTLTGRPPFDIVTNTLKNVVLADYEMPSFLISAKDLIHQLRRNPADRL 254  
DB 205 WSLGCMFYTLTGRPPFDIVTNTLKNVVLADYEMPSFLISAKDLIHQLRRNPADRL 254  
QY 255 SLSSVLDHPFMSRNSSTKSLDGTVEDSIDSGHATISTATSSST-----SIG 304  
DB 265 SLDDQLRHDFPK-----GYTPDRL-----PSSCTVVDLTPNPABSLPAKVT 310  
QY 305 SLFDRK 311  
DB 311 SLFGRK 317

## RESULT 9

US-10-755-889-4  
Sequence 4, Application US/10755889  
Publication No. US20040171823A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB  
FILE REFERENCE: D0284 NP  
CURRENT APPLICATION NUMBER: US/10/755,889  
CURRENT FILING DATE: 2004-01-13  
PRIOR APPLICATION NUMBER: U.S. 60/440,068  
PRIOR FILING DATE: 2003-01-14  
PRIOR APPLICATION NUMBER: U.S. 60/469,757  
NUMBER OF SEQ ID NOS: 823  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 4  
LENGTH: 646  
TYPE: PR  
ORGANISM: Homo sapiens  
US-10-755-889-4

Query Match 11.6%; Score 590.5; DB 16; Length 646;

Best Local Similarity 37.5%; Pred. No. 7.7e-32;  
Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;

QY 15 GNLLGSGFAGVYRABSIHTGLEVAIKMIDKKAMYGAVORVONEVKIHCQKPSILE 74  
DB 26 GNLLGSGFAGVYRABSIHTGLEVAIKMIDKKAMYGAVORVONEVKIHCQKPSILE 74

[illegible]

```

RESULT 10
US-10-026-021-4
Sequence 4, Application US/10026021
Publication No. US20030027756A1
GENERAL INFORMATION:
APPLICANT: Hitoshi, Yasumichi
APPLICANT: Demo, Susan
APPLICANT: Jenkins, Yonchu
APPLICANT: Rigel Pharmaceuticals, Inc.
TITLE OF INVENTION: SAR: Modulation of Cellular Proliferation for
TITLE OF INVENTION: Treatment of Cancer
FILE REFERENCE: 021044-001210US
CURRENT APPLICATION NUMBER: US/10/026,021
CURRENT FILING DATE: 2002-06-25
PRIOR APPLICATION NUMBER: US 60/309,632
PRIOR FILING DATE: 2001-08-01
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 373
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (1)..(373)
OTHER INFORMATION: human FNK mitotic kinase kinase domain
US-10-026-021-4

```

Query Match	11.6%	Score 589.5	DB 14	Length 373
Best Local Similarity	37.5%	Pred. No. 4.28-32		
Matches 115; Conservative	61	Mismatches 106	Indels 25	Gaps 4

[illegible]

Db 304 SLDQIRHDFYTK-----GYTPDL-----PISCTVBDLTPPNPARSLFAVTK 349

Qy 305 SLFDKRR 311  
|||:::  
Db 350 SLFVRKK 356

```

RESULT 11
US-10-620-052A-43
; Sequence 43, Application US/10620052A
; Publication No. US2004012678A1
; GENERAL INFORMATION:
; APPLICANT: Hitoshi, Yasumichi
; APPLICANT: Jemkins, Yomchu
; APPLICANT: Markovtsov, Vadim
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: Modulators of Cellular Proliferation
; FILE REFERENCE: 021044-004010US
; CURRENT APPLICATION NUMBER: US/10/620, 052A
; CURRENT FILING DATE: 2003-07-14
; PRIOR APPLICATION NUMBER: US 60/395,443
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 253
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: cytokine-inducible kinase (CNK) serine threonine
; OTHER INFORMATION: kinase, proliferation-related kinase (PRK),
; OTHER INFORMATION: polo-like kinase 3 (PLK3)
; US-10-620-052A-43

```

Query March	11.4%	Score 580.5	DB 16	Length 253
Best Local Similarity	42.4%	Pred. No. 1e-31		
Matches 106, Conservative	51	Mismatches	92	Gaps 1

```

QY      15 GNLLKKGSAGAYRRAESIHTLEVAIKMIDKKAMVQVONEVKIHCOLNHSILE 74
Db      4 GNLLKGGSFARCYEATDTJETSAAVAVKTIPOSRVAKQHQEKILNELIELHRDLOHRIIVR 63
QY      75 LNVYEDSNVYVYLVEMHCHNEBNRYLKNRXPFSSEBAHHFHHQITGMLYLHSHGILH 134
Db      64 FSHHEEDADNNTYIFLELCRSKSLAHNIWPAR-HTLLEBEVRVYLLQJLISGKLYLHQRGILH 122
QY      135 RDLTSLNNLLTFNNNNIKIADFGALATOLKMPHBEKHVYTCGTPNTYISPEIATRSASHGLESVD 194
Db      123 RDLKLGKGFITENMELKIGDFGLAARLEPPRQRKGTICGTPNYAARVLLRQGHGPRADV 182
QY      195 WSLGCMFYTLILGRPEPDITVKNTLKNVLADYEMPSLEAKDILIHOLLRRNPADRL 254
Db      183 WSLGCMVMTYLLCGSPPEFTADLKETVRCIKQVHTYLLPASLSLPAHQLLAAILRLASPRDRP 242
QY      255 SLSSVLDHNP 264
Db      243 SIDQILRHDF 252

RESULT 12
US-10-618-581-15
; Sequence 15, Application US/10618581
; Publication No. US20040077524A1
GENERAL INFORMATION:
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
TITLE OF INVENTION: METHOD FOR SELECTIVELY INHIBITTING FUNGAL GROWTH
FILE REFERENCE: 034536/0323
CURRENT APPLICATION NUMBER: US/10/618, 581
CURRENT FILING DATE: 2003-07-15
PRIORITY APPLICATION NUMBER: 60/395,624
PRIORITY FILING DATE: 2002-07-15
NUMBER OF SEQ ID NOS: 42

```

SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 15  
 ; LENGTH: 1462  
 ; TYPE: PRT  
 ; ORGANISM: Candida albicans  
 US-10-618-581-15

Query Match 11.2%; Score 568.5; DB 16; Length 1462;  
 Best Local Similarity 26.5%; Pred. No. 7.8e-30;  
 Matches 189; Conservative 107; Mismatches 255; Indels 163; Gaps 21;

```

QY      8 KIEDPKVGNLKGSGFAGYVRAESIHITGLEVAIKMIDKAMK-----AGNV 54
DB      61 KIGPKLGTGRTGSGTGRVRLAKNTTGGDLAAVKIYPKSNPKLEMPKXSKEDATRLP 120
QY      55 QRVQNEVKIKHCOLKHPISIELNYNPFDSNVYVLVLEMGCHGEMNRKYLKRVKPFSENEK 114
DB      121 YGIREBIIIMKLIISPINIMGLYDVWENKQDLYLLEIYIGGELFDYILIRGK-LDEYEAI 179
QY      115 HPMHQITGMLYHSHGILHRDLTSLNLLTTRNMKIKADFGLATQMKPHEKHYTLGCT 174
DB      180 NFKQIINGINYLHGFNICHRLDKPENLLDFKNIKIKADFGMAA-LEVYKELLETSCS 238
QY      175 PNYISPEI-ATRSAGLESDVMSLGCIFYTLIGRPFPDITVKNLANKVVLADYEMPEF 233
DB      239 PNYASPEIYAGKNYHAGPDSIWSGIIILFALLTGHLPFDENIRKLLKLVQSGKFMPE 298
QY      234 LIEAKDLIHQLLRNPADRLSLSVLDHPFMSR-----NSYKSKDLGTYE---DS 282
DB      229 LSPFAKDLITKMLKVPREPRITIDALITPLAKYDEPTVSYSTTLLDINSINIKQIES 358
QY      263 IDSGHATITAT-----ASSSTISGSLPDKRILLIGPLPKYMTVPFKNK-----S 330
DB      359 VKIDKEIILKNLSVLFHNCDEKTLIS-----RLSPNRCPEKMFYLLMKTRNEHLSN 411
QY      331 STDFSSGCD-----GNSFYTOMGNOETSN 355
DB      412 SNFSFSSNDVDARSIPRSTSVKTYTVDHANGEKHTYKLIQSSSIYSNLSLKKSS 471
QY      356 GGRVYIQDAEER-----HSRYLRAYSSDRGTSNSQOAKTYTME 397
DB      472 AKGNVLSNTNRPNTPKQPSASSSFNKKALHSK--TQIYASRRNASRSLKSNSTOR 529
QY      398 RCHSAEMLSVKR-----RSGGEMER-----YSPDNNANFNF--PK 434
DB      530 NGNNASVTSVANKIPEITGATVLIQIPSMAMNRGDEQONTKKNLTGTFPGKSLNFIOLIC 589
QY      435 EKTSSSGSFERRPDNNQALSNHLCPGKTFPPFADPPQETVQQ--WF-----480
DB      590 EAVFENDKENSKEPVSKTPVSO--LPPPPPIETPTSRNSVYKGTWLSARERELAQ 647
QY      481 -----GNLQINAHLRKTEYD-----SISPNRDFQGHPI--QKDTSKNAWTDYV 524
DB      648 VQRNREARENKLAEBELARKELEQEKRIABEKKRLQEQERELDEKQKLEOKKALERKL 707
QY      525 KKNSDASD-----NAHSVQOQNTKMTALHAKREIITQOECVPSDPLSEBS 571
DB      708 QKQSHADEGLFASNRRSVTDMAPSSGMSLDPRAHWVSRAITGSPMLSSS 761

```

RESULT 13  
 US-10-026-021-5

Sequence 5; Application US/10026021  
 Publication No. US20030027756A1  
 GENERAL INFORMATION:  
 APPLICANT: Hitooshi, Yasumichi  
 APPLICANT: Demo, Susan  
 APPLICANT: Jenkins, Yonchu  
 APPLICANT: Rigel Pharmaceuticals, Inc.  
 TITLE OF INVENTION: SAK: Modulation of Cellular Proliferation for  
 FILE REFERENCE: 021044-001210US  
 CURRENT APPLICATION NUMBER: US/10/026, 021

CURRENT FILING DATE: 2002-06-25  
 PRIOR APPLICATION NUMBER: US 60/309, 632  
 PRIOR FILING DATE: 2001-08-01  
 NUMBER OF SEQ ID NOS: 8  
 SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 5  
 ; LENGTH: 400  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 NAME/KEY: DOMAIN  
 LOCATION: (1)..(400)  
 ; OTHER INFORMATION: human SNK mitotic kinase kinase domain  
 US-10-026-021-5

Query Match 11.0%; Score 560.5; DB 14; Length 400;  
 Best Local Similarity 41.6%; Pred. No. 4.6e-30;  
 Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

```

QY      15 GNILKGSPAGYVRAESIHITGLEVAIKMIDKAMVAGVQVQNEVKIKHCOLKHPISILE 74
DB      85 GKVLGKGFPAKCYEMTDLTNKNVYAKIIPHSRVAKPHQRKIDKEIELRIHHGAVQ 144
QY      75 LNYFEDSNVYVLVLEMGCHGEMNRKYLKRVKPFSENEARHFMHQIITGMLYHSHGILH 134
DB      145 FYHFEDEKENIYILLEYGSGRSMAILKAR-KVLTPEVRYVLRQIVSGIKYLEQELIH 203
QY      135 RDLTSLNLLTTRNMKIKADFGLATQMKPHEKHYTLCTPNYISPEIATRSAGLESDV 194
DB      204 RDLKGNFPIINAMEIKVDFGLARLEPLERHRRRTCTPNYLSSEVLNKOQGESDI 263
QY      195 WSIGCMFYTLIGRPFPDITVKNLANKVVLADYEMPSFLIEAKDLIHQLLRNPADRL 254
DB      264 WDLGCVMTMLGRPFETITNLKETYRCIRBARTYTPSSILPAGHLIASMLSKNEDRP 323
QY      255 SLSSVLDHPF 264
DB      324 SLDDIIRHDF 333

```

RESULT 14  
 US-09-771-161A-249  
 Sequence 249; Application US/09771161A  
 Patent No. US20020110811A1  
 GENERAL INFORMATION:  
 APPLICANT: LEVINE, et al.  
 TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
 FILE REFERENCE: 802620-2005.1  
 CURRENT APPLICATION NUMBER: US/09/771,161A  
 CURRENT FILING DATE: 2001-01-26  
 PRIOR APPLICATION NUMBER: 09/724,676  
 PRIOR FILING DATE: 2000-11-28  
 PRIOR APPLICATION NUMBER: 136776  
 PRIOR FILING DATE: 2000-06-15  
 PRIOR APPLICATION NUMBER: 135619  
 PRIOR FILING DATE: 2000-04-12  
 NUMBER OF SEQ ID NOS: 273  
 ; SOFTWARE: Patentin version 3.0  
 ; SEQ ID NO 249  
 ; LENGTH: 685  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-771-161A-249

Query Match 11.0%; Score 560.5; DB 9; Length 685;  
 Best Local Similarity 41.6%; Pred. No. 9.7e-30;  
 Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

```

QY      15 GNILKGSPAGYVRAESIHITGLEVAIKMIDKAMVAGVQVQNEVKIKHCOLKHPISILE 74
DB      85 GKVLGKGFPAKCYEMTDLTNKNVYAKIIPHSRVAKPHQRKIDKEIELRIHHGAVQ 144
QY      75 LNYFEDSNVYVLVLEMGCHGEMNRKYLKRVKPFSENEARHFMHQIITGMLYHSHGILH 134

```

```
Db      145 FHYHFEDEKENIYIYLEYCSRRSMHILKAR-KVLTEPEVRYYLRQIVSGIKYLIHQEILH 203
QY      135 RDLTSLNLLITNNMIKIDFGIATQOLKMPHEKHYTLGTPNYISPEIATRSAGHLES DY 194
Db      204 RDLKLGNEFINAMELKVDGFGIARLEPLEHERRRRTICGTPNYLSPEVLNKGHGESDI 263
QY      195 WSLGCMFYTLILGRPPFDITVKNTLKVVLADYEMPSFLSIEAKDLIHQLLRNPADRL 254
Db      264 WALGCVMTYMLLGRPPFETTNKETYRCIRBARYTMPSSILAPAKHLIASMLSKNPEDRP 323
QY      255 SLSSVLDHPF 264
Db      324 SLDDIIRHDF 333
```

## RESULT 15

```
US-09-771-161A-250
; Sequence 250, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 250
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-250
```

Query Match 11.0%; Score 560.5; DB 9; Length 685;

Best Local Similarity 41.6%; Pred. No. 9,7e-30; Mismatches 99; Indels 1; Gaps 1;

```
Db      15 GNILGKSPAGVYRABSIHITGLEVAIKMIDKAMYKAGVORVONEVKIHCQLKPSILE 74
QY      85 GNVLGKSGPAKCYEMTDLTNNKVYAAKITPHSRVAKPHOREKIDKEIHLRIHKKAVYQ 144
Db      75 LYNVFEEDSNVYVLVLEMCNNGEMNRYLKNRVKPFSENEARHFMHQIITGMLYIASHGILH 134
QY      145 FHYHFEDEKENIYIYLEYCSRRSMHILKAR-KVLTEPEVRYYLRQIVSGIKYLIHQEILH 203
Db      135 RDLTSLNLLITNNMIKIDFGIATQOLKMPHEKHYTLGTPNYISPEIATRSAGHLES DY 194
QY      204 RDLKLGNEFINAMELKVDGFGIARLEPLEHERRRRTICGTPNYLSPEVLNKGHGESDI 263
Db      195 WSLGCMFYTLILGRPPFDITVKNTLKVVLADYEMPSFLSIEAKDLIHQLLRNPADRL 254
QY      264 WALGCVMTYMLLGRPPFETTNKETYRCIRBARYTMPSSILAPAKHLIASMLSKNPEDRP 323
Db      255 SLSSVLDHPF 264
QY      324 SLDDIIRHDF 333
```

Search completed: September 23, 2004, 21:04:13  
Job time : 72 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 20:58:26 ; Search time 23 Seconds

(without alignments)  
4056.773 Million cell updates/sec

Title: US-10-026-021-2

Perfect score: 5078

Sequence: 1 MATCGEKIEDPKVGNLTKGK.....KLQCLSSILMFSPNTPNPH 970

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3927.5	77.3	925	2	A55748	protein kinase (EC
2	1921.5	37.8	465	2	B55748	protein kinase (EC
3	593	11.7	631	2	A57286	probable serine/th
4	591.5	11.6	576	2	S22127	protein kinase pol
5	561.5	11.1	682	2	A44493	serum-inducible k
6	554.5	10.9	407	2	S52343	p46Bg265 protein-
7	548	10.8	389	2	S52242	protein kinase (EC
8	545	10.7	603	2	A47545	protein kinase (EC
9	545	10.7	603	2	A54596	protein kinase - m
10	544.5	10.7	1246	2	G89287	protein kinase H3B23.1
11	536	10.6	603	2	S34130	serine/threonine-s
12	527	10.4	461	2	T14822	probable serine/th
13	527	10.4	1192	2	T18611	probable serine/th
14	526.5	10.4	1518	2	S37928	probable purine nu
15	514.5	10.1	683	2	T38254	serine/threonine-s
16	511.5	10.1	887	2	T20941	hypothetical prote
17	504.5	9.9	403	2	JCS974	autora-related kin
18	501	9.9	648	2	T43337	polo-like kinase-1
19	500	9.8	547	2	S44841	KO6H7.1 protein -
20	498	9.8	615	2	T29223	hypothetical prote
21	496	9.8	512	1	JC1446	serine/threonine-s
22	493.5	9.7	1142	2	S59359	GIN4 protein - yea
23	491	9.7	395	2	JCS975	autora-related kin
24	490	9.6	705	2	A48144	protein kinase CDC
25	489	9.6	441	2	B85362	hypothetical prote
26	487.5	9.6	384	2	T41298	ser/chr protein k
27	486.5	9.6	502	2	T02306	probable protein k
28	486	9.6	915	2	S74283	probable protein k
29	484	9.5	512	2	T52633	serine/threonine-s

30	480.5	9.5	345	2	JC4665	protein kinase (EC
31	480	9.5	480	2	A86427	probable serine/th
32	479.5	9.4	504	2	T10449	probable serine/th
33	479	9.4	520	2	G86414	probable protein k
34	478.5	9.4	745	2	G01025	serine/threonine p
35	477.5	9.4	445	2	T09609	serine/threonine-s
36	476	9.4	445	2	T09609	serine/threonine-s
37	473	9.3	489	2	T04862	probable serine/th
38	470.5	9.3	456	2	T46478	serine/threonine p
39	469	9.2	367	2	S47923	probable protein k
40	469	9.2	453	2	G86141	protein T25K16.13
41	468	9.2	1081	2	S51899	probable protein k
42	465	9.2	592	2	T43402	probable protein k
43	464.5	9.1	473	1	S59941	serine/threonine-s
44	464	9.1	1558	2	T29253	hypothetical prote
45	462.5	9.1	633	1	A26030	serine/threonine-s

#### ALIGNMENTS

##### RESULT 1

A55748 protein kinase (EC 2.7.1.37) Sak-a - mouse

C/Species: Mus musculus (house mouse)

C/Date: 02-Jun-1995 #sequence\_revision 02-Jun-1995 #text\_change 24-Sep-1999

C/Accession: A55748

R/Fode, C.; Motro, B.; Yousefi, S.; Helferman, M.; Dennis, J.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 6388-6392, 1994

A/Title: Sak, a murine protein-serine/threonine kinase that is related to the Drosophila

A/Accession number: A55748; MUID:94294387; PMID:8022793

A/Accession: A55748

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-925 <FOD>

A/Cross-references: GB:129479; NID:548769; PIDN:AAC37648.1; PID:5487670

A/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C/Keywords: ATP; phosphotransferase

F/10-265/Domain: protein kinase homology <KIN>

F/18-26/Region: protein kinase ATP-binding motif

Query Match 77.3%; Score 3927.5; DB 2; Length 925;  
Best Local Similarity 78.6%; Pred. No. 3.8e-146;  
Matches 763; Conservative 76; Mismatches 83; Indels 49; Gaps 9;

QY	1	MATCGEKIEDPKVGNLTKGKSPAGYTRASHTGLVAIKMIDKAMKAGVQRYONE	60
DB	1	MAACIGERIEDPKVGNLTKGSPAGYTRASHTGLVAIKMIDKAMKAGVQRYONE	60
QY	61	VKTHCOLKRPSTILELVNPFEDSNVYVLYLVMCHNGEMNRYLKNRVKPFSENEARHPMHQI	120
DB	61	VKTHCOLKRPSTILELVNPFEDSNVYVLYLVMCHNGEMNRYLKNRVKPFSENEARHPMHQI	120
QY	121	ITGMVLIHSHGILHRDLTSLNLLTTRNNIKIADFGIATQLKMPHEKHVTLCTPYISF	180
DB	121	ITGMVLIHSHGILHRDLTSLNLLTTRNNIKIADFGIATQLKMPHEKHVTLCTPYISF	180
QY	181	ELATRSAGLSDVMSLGCFFTLIGRPPEPTDTYKATLANKVVLADYEMPSFLSLEAD	240
DB	181	ELATRSAGLSDVMSLGCFFTLIGRPPEPTDTYKATLANKVVLADYEMPSFLSLEAD	240
QY	241	LHOLLRRNPADRLSSVLDHPFMSRNSSTKSDIGTVEDSIDSGHATSTLITSSST	300
DB	241	LHOLLRRNPADRLSSVLDHPFMSRNSSTKSDIGTVEDSIDSGHATSTLITSSST	300
QY	301	SISGSLFDRRLIGQPLPNKMTVPKPKNSSTDFSSGSGDGSFFYQWGN--OETNSGRG	358
DB	301	SISGSLFDRRLIGQPLPNKMTVPKPKNSSTDFSSGSGDGSFFYQWGN--OETNSGRG	358
QY	359	RVYQDAERPHSRVYLRAYSSDSGTSNSQSQAKYTMERCHSAEVLSTYSKSGGSENEE	418
DB	359	RVYQDAERPHSRVYLRAYSSDSGTSNSQSQAKYTMERCHSAEVLSTYSKSGGSENEE	418

```
Qy 419 RYSPPTDNNANIFNPFKEKTSSSSGSFERRPDDNNQALSNHLCPGKTPFPFPADPTPOTETVQO 478
Db 411 -----LDERQHSNNHCIGKTFPFPADPTPQEMEWQO 442
Qy 479 MFGNQLMAHLKRTETYSISPNRDPQGHPIQKDTSNAMTDTKVKKNSDASDNASVK 538
Db 443 MFGNQLMAHLGENTNEHTVSPNRDPQDYPDIQ-DITLNAWTDPRASKNADTSANVAVK 501
Qy 539 QONTKMTATLHSPKEIIOQECVFGSDPLSEOSKTRGMEPPWGMONTLSITSPVLAHR 598
Db 502 QLSAMKVSAAHHNREVPQRP--GLHPSIOSKRSWESTLGYQKPLSISTPLAHR 559
Qy 599 LKPIRQKTKKAVSILDSEEVCEVLVREYASQEVYKVELQISSDGMTITITYPNGRGRFP 658
Db 560 LKPIRQKTKKAVSILDSEEVCEVLVREYASQEVYKVELQISSDGMTITITYPNGRGRFP 619
Qy 659 LADPPSPPTDNISSYSPNLPBKTKWRYQVARSFVCLVRSKSPITTYTTRAKLIMENS 718
Db 620 LADPPPLPTDNISYSPNLPBKTKWRYQVARSFVCLVRSKSPITTYTTRAKLIMENS 679
Qy 719 PGADPEWVFYDGVKIKHKTEDFIQVIEKTKGKSYTLKSEEVNSLKEBITKMTMDHANEGRH 778
Db 680 PGADPEWVFYDGVKIKHKTEDFIQVIEKTKGKSYTLKSEEVNSLKEBITKMTMDHANEGRH 739
Qy 779 CLAMESTISEERKTRAPPPPIIIGRKPGSTSPKALSPPPSVDSNYPTRDRASFNRWV 838
Db 740 CLAMESTISEERKTRAPPPPIIIGRKPGSTSPKALSPPPSVDSNYPTRDRASFNRWV 798
Qy 839 MHSNAPSPTQAPILNPSWVTEGLGLTTASGTDISSSLDCLPKSAQLIKSPVNVNWM 898
Db 799 VNSAAPPFGTQSPGSPVTVYEGIGHRTATAGTVSSS-----LPKSAQLIKSPVNVNWM 853
Qy 899 ATQLTSGAVVQVRNDSQVLVQAGVSSISYSPNGQTRRGENEKLPDYIKQRLQCLSSI 958
Db 854 ATQLTSGAVVQVRNDSQVLVQAGVSSISYSPNGQTRRGENEKLPDYIKQRLQCLSSI 913
Qy 959 LLMFSNPTPNF 969
Db 914 LLMFSNPTPNF 924
```

```
RESULT 2
B55748
protein kinase (EC 2.7.1.37) Sak-b - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 19-Dec-1997
C:Accession: B55748
R:Code, C.; Mottu, B.; Yousefi, S.; Hefferman, M.; Dennis, J.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 6388-6392, 1994
A:Title: Sak, a murine protein-serine/threonine kinase that is related to the Drosophila
A:Reference number: A55748; MUID:94294387; PMID:8022793
A:Accession: B55748
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-465 <F0D>
A:Cross-references: GB:I29480
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; phosphotransferase
F:10-265/Domain: protein kinase homology <KIN>
F:18-26/Region: protein kinase ATP-binding motif
```

```
Query Match 37.8%; Score 1921.5; DB 2; Length 465;
Best Local Similarity 84.5%; Pred. No. 2.5e-68;
Matches 370; Conservative 29; Mismatches 34; Indels 5; Gaps 4;
```

1 MATCHGKTEIDPKVGNLLGKSPAGVYRAESIHTELEVAIKMTDKKAMYKAGVQVQVONE 60  
1 MAACIGETIEIDPKVGNLLGKSPAGVYRAESIHTELEVAIKMTDKKAMYKAGVQVQVONE 60

61 VKIHQQLHSPSILLEYNFEDSNVYVLYLEMCHNGEMNRVYKRVKPSSEAEAHFMHQI 120  
61 VKIHQQLHSPSILLEYNFEDSNVYVLYLEMCHNGEMNRVYKRVKPSSEAEAHFMHQI 120

```
Qy 121 ITGMLYHSHGILHRDLTSLNLTTRNNIKIADFGIATOLKMPHEKHYTLGTPYISF 180
Db 121 ITGMLYHSHGILHRDLTSLNLTTRNNIKIADFGIATOLKMPHEKHYTLGTPYISF 180
Qy 181 EINTRSAGHLESPVWLSGCMFYTLILGRPPPTDVTYKNTLKVYLADYEMPSFLISIAEKD 240
Db 181 EINTRSAGHLESPVWLSGCMFYTLILGRPPPTDVTYKNTLKVYLADYEMPSFLISIAEKD 240
Qy 241 LIHQLLRRNPADRLSLSSVLDHPFMSRNSSTSKDGLGYEDSIDSGHATISPAITASST 300
Db 241 LIHQLLRRNPADRLSLSSVLDHPFMSRNSSTSKDGLGYEDSIDSGHATISPAITASST 300
Qy 301 SISGSLPDKRILLIGQPLPNKATVPKKNKSTDPSSSGDGSFYTQWGN--QETSNSGRG 358
Db 301 SISGSLD--RRLLVQQLPNKTIIVFOKKNSSDF--SSGDGSFYTQWGNPQEAANSRGRG 358
Qy 359 RYIODEERPHSRVYLRRAVSSDSRSTNSOSQAKTYMTERCHSAEMLSVSKSGGGENBE 418
Db 359 RYIODEERPHSRVYLRRAVSSDSRSPSN--QSRKTIYVERCHSVEMLSKPRSLDENQHS 417
Qy 419 RYSPPTDNNANIFNPFKEK 436
Db 418 RYSPTKSNVNVLTSLNTK 435
```

```
RESULT 3
A57286
probable serine/threonine protein kinase (EC 2.7.1.-) fnk - mouse
C:Species: Mus musculus (house mouse)
C:Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 10-Sep-1997
C:Accession: A57286
R:Donohue, P.J.; Alberts, G.P.; Guo, Y.; Winkles, J.A.
J. Biol. Chem. 270, 10351-10357, 1995
A:Title: Identification by targeted differential display of an immediate early gene encoc
A:Reference number: A57286; MUID:95247749; PMID:7730342
A:Accession: A57286
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-631 <DON>
A:Cross-references: GB:U21392; GB:U22434
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; phosphotransferase
F:61-315/Domain: protein kinase homology <KIN>
```

```
Query Match 11.7%; Score 593; DB 2; Length 631;
Best Local Similarity 33.1%; Pred. No. 1.9e-16;
Matches 122; Conservative 64; Mismatches 111; Indels 72; Gaps 4;
```

15 GNLLGKSPAGVYRAESIHTELEVAIKMTDKKAMYKAGVQVQVQVNTIKHCKHPSILE 74  
66 GRLLGKGFARCYEATDTESGIAYAVVLPQSRVAKPHQREKILNLELHRDLQHGHYR 125

75 LVNPFEDSNVYVLYLEMCHNGEMNRVYKRVKPSSEAEAHFMHQIITGMLYHSHGILH 134  
126 FSHHFEADNDIYVLELCRSKRSIAHWKAR-HTLLEPEVRYVLRQILSGIKYLHQGHYR 184

135 RDLTTLNLTTRNNIKIADFGIATOLKMPHEKHYTLGTPYISPAITRSAGHLESPV 194  
185 RDLKLGNFPTDMELKQVDFGLAARLEPPEQKKTICGTYYVAEVLIRQCHGEADV 244

195 WSLGCMFYTLILGRPPPTDVTYKNTLKVYLADYEMPSFLISIAEKDLIHQLLRRNPADRL 254  
245 WSLGCMFYTLILGRPPPTDVTYKNTLKVYLADYEMPSFLISIAEKDLIHQLLRRNPADRL 304

255 SLSSVLDHPFMSR-----NSTSKDLG 277  
305 SIQILRHFPFTKGYTPDLTPVSSCYTTPDLTPNPANSLFAKVTKSLFGRRKNKNKHS 364

278 TWEDSID-----SGHATISAITASSSTISGSLFDKRRLLIGQLPNKMTVPFKNRS 331  
365 EPQDNVSCIAPVVSGQAPASLLETAAEDSSPRGLT----- 399

332 TDPSSSGDG 340



Db 400 ----ASSGDCG 405

## RESULT 4

S22127  
protein kinase polo (EC 2.7.1.-) - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 24-Sep-1999  
C:Accession: S22127  
R:Sunkel, C.B.  
submitted to the EMBL Data Library, November 1991  
A:Reference number: S22127  
A:Accession: S22127  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-576 <SUN>  
A:Cross-references: EMBL:X63361; NID:g8355; PIDN:CAA44963.1; PID:g8356  
C:Genetics:  
A:Gene: PLYBase:polo  
A:Cross-references: FLYBase:FBgn0003124  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP; phosphotransferase  
F:23-27/Domain: protein kinase homology <KIN>

Query Match 11.6%; Score 591.5; DB 2; Length 576;  
Best Local Similarity 33.8%; Pred. No. 26-16; Indels 71; Gaps 9;  
Matches 135; Conservative 71; Mismatches 122;

2 ATCTGKED-----FKVGNLKGKSPAGVYRAESITHTGLEVALIKMDKAMYGKAMVQ 55  
9 STDIPRLVDINQKRTYKMRPFKGGKFAKCYEIIIDVETDDVAGKIVKGMKINQKE 68  
56 RVQNEVYKICQLKPSILELYNTEFDSNTYVYLEKCHNGENRKYLNRYKPESENEAR 115  
69 KTAQETITIRSLNHPVIFVFNFEFSQNIYVLELCKKSMELKRR-KSTTEPCY 127  
116 FMHOITITGLVYSHSHILRDILTLNMLLITRNNTIADFGIATOLKMHKHYTLGTP 175  
128 YVQIITIGKYLHNDRIHRDLKGLNLFNDLHKGIGLATRIEYEGEKTKTCGTA 187  
176 NTSPETIRSAHGLESDVMSLGMFYTLIGRPFDVTYKNTLKNVLADEMPFLS 235  
188 NYIAPBILTKGHSFVDLMSICWYTLINQPPETITIKDYTKIKCEKRVSYLR 247  
236 IBAKDLIHLRLRNPRDLRLSSVLDHPFMS-----RNSSTYSKDLG--TVEDSI 283  
248 KPAADVYIAMLQPNPESRAIGQLNFEFLKSKVPMFLPSSCLTAPRIGSNDTIEDSM 307  
284 D-----SG-----HATISRAITASSSTISGSLFDRRLILQPLPNKMT 323  
308 HKRPLEMNGIRPDDTRLESTFLKALHAIITASQVC----- 345  
324 VEPKNSSTDFSSGDSNGFYTQWNGQETSNSGRGVIO 362  
346 -----RHSEDYNS--DIESLYOQLTN---LINSKRILO 374

## RESULT 5

A44493  
serum-inducible kinase - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 10-Sep-1997  
C:Accession: A44493  
R:Simmons, D.L.; Neel, B.G.; Stevens, R.; Ewelt, G.; Erikson, R.L.  
Mol. Cell. Biol. 12, 4164-4169, 1992  
A:Title: Identification of an early-growth-response gene encoding a novel putative protein kinase  
A:Reference number: A44493; MUID:92375085; PMID:1508211  
A:Accession: A44493  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-682 <SIM>  
A:Experimental source: F-2 cells

A>Note: sequence extracted from NCBI backbone (NCBIN:111721, NCBIPI:111722)  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP  
F:77-331/Domain: protein kinase homology <KIN>

Query Match 11.1%; Score 561.5; DB 2; Length 682;  
Best Local Similarity 41.6%; Pred. No. 3-5e-15;  
Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

15 GNILKGKSPAGVYRAESITHTGLEVALIKMDKAMYGKAMVQVQNEVYKICQLKPSILE 74  
82 GRVLDGKSGFAKCYEMTDLTNKNVYAKIIPHSRVAPPHOREKIDKIELHRLHNGVQ 141  
75 LYNTEFDSNTYVYLEKCHNGENRKYLNRYKPESENEARFMHOITITGLVYSHSHI 134  
142 FYHVEDKENTITILEYCCRRSMHILKAR-KVLTPEVRYVLRQVSGIKLHGEIILH 200  
135 RDLTSLNMLITRNNTIKIADFGIATOLKMHKHYTLGTPNYISPEIATRSAGLESVD 194  
201 RDLKGNFPIINEMELKVDGFLAARLEPLERHRRITCGTPVYLSPEVLNKGHGESDI 260  
195 WSLGCMFYTLIGRPFDVTYKNTLKNVLADEMPFLSBAKDLIHLRLRNPRDL 254  
261 WALGCMYTMILGRPEFTTNLKEHYRCIRARYTTPSSILAPAKLILASMLSKNEDRP 320  
255 SLSSVLDHPF 264  
321 SLDDIIRHDF 330

## RESULT 6

S52243  
p45g265 protein - African clawed frog

C:Species: Xenopus laevis (African clawed frog)  
C:Date: 07-May-1995 #sequence\_revision 03-Aug-1995 #text\_change 24-Sep-1999  
C:Accession: S52243; S34642; I51695  
R:Koghi, C.; le Guellec, R.; Paris, J.; Couturier, A.; Philippe, M.  
submitted to the EMBL Data Library, October 1992  
A:Description: Egt, selected by differential screening encodes a new Xenopus protein kinase  
A:Reference number: S52242  
A:Accession: S52242

A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-407 <ROG>  
A:Cross-references: EMBL:217207; NID:g609281; PIDN:CAA78915.1; PID:g609282  
R:Bouvet, P.; Omillii, F.; Ariot-Bonnemains, Y.; Legagneux, V.; Koghi, C.; Bassez, T.; Oebc  
submitted to the EMBL Data Library, June 1993  
A:Description: Targeted deadenylation of specific mRNAs in Xenopus embryos by a mechanical  
A:Reference number: S34642  
A:Accession: S34642

A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 328-407 <BOU>  
A:Cross-references: EMBL:224453; NID:g394756; PIDN:CAA80826.1; PID:g394757  
R:Bouvet, P.; Omillii, F.; Ariot-Bonnemains, Y.; Legagneux, V.; Koghi, C.; Bassez, T.; Oebc  
Mol. Cell. Biol. 14, 1893-1900, 1994  
A:Title: The deadenylation conferred by the 3' untranslated region of a developmentally c  
A:Reference number: I51695; MUID:9415861; PMID:8114721  
A:Accession: I51695

A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 328-407 <BO2>  
A:Cross-references: EMBL:224453; NID:g394756; PIDN:CAA80826.1; PID:g394757  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP; phosphotransferase; protein kinase  
F:138-390/Domain: protein kinase homology <KIN>  
F:146-154/Region: protein kinase ATP-binding motif

Query Match 10.9%; Score 554.5; DB 2; Length 407;  
Best Local Similarity 39.7%; Pred. No. 3-8e-15;  
Matches 106; Conservative 60; Mismatches 98; Indels 3; Gaps 2;

9 IEDFKVGNLKGKSPAGVYRAESITHTGLEVALIKMDKAMYGKAMVQVQNEVYKICQLK 68

```

Db      137 LEDPEIGRPLGKGRGVNYLAREERKRIALLKTLFPGSOLKAEVGHQLRREVAIQLSHLR 196
      69 HPSILLEYNPFEDSNVYVLYLEKCHNGEMNRYLKORVYKPSSENAHFMQIITGMULYH 128
      197 HENILRLYGYHDSRYYLLIDYAPGELFRELQKCTR-FDDQSRNAYIKQALBALYCH 255
Qy      129 SHGILHDLTSLNLLTFNNMIKTLADFGLATQLMKPHKHYTLGGTPNYISPEIATRSAN 188
      256 SKKVIHRDIPKRNLLSGNSGLKIADFGMS--VHAPSSRRTLLCGTLDYLPPEMIEGRMH 313
Db      189 GLESDVWSLGCMPFTLLIGRPEDPTDVKTANKVNLADYEMPSFLSIEAKDLIHQILRR 248
      314 DETYDLSLGLVLCYEFVVGKPPFETDTHQETVRIISKVEFOYPPYVSEBARDLVSKLKH 373
Qy      249 NPADRLSLSSVLDHPFMSRNSSTKSD 275
      374 NPNHRLPLKGVLEHPIWIKNSQLKKKD 400

```

```

RESULT 7
552242
Protein kinase (EC 2.7.1.-) p46Xlg22 - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 18-Jun-1999
C/Accession: S52242
R/Rogni, C.; Le Guellec, R.; Paris, J.; Coucunier, A.; Philippe, M.
Submitted to the EMBL Data Library, October 1992
A/Description: Bg2, selected by differential screening encodes a new Xenopus protein kin
A/Reference number: S52242
A/Accession: S52242
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-389 <RGS>
A/Cross-references: EMBL:Z17206; NID:9609279; PIDN:CAA78914.1; PID:9609280
C/Superfamily: kinase-related transforming protein; protein kinase homology
C/Keywords: ATP; phosphotransferase
F:119-371/Domain: protein kinase homology <KIN>
F:127-135/Region: protein kinase ATP-binding motif

```

```

Query Match      10.8%; Score 548; DB 2; Length 389;
Best Local Similarity 39.9%; Pred. No. 6.5e-15;
Matches 107; Conservative 60; Mismatches 97; Indels 4; Gaps 3;

Qy      9 IDPKVGNLKGKSPAGVYRAESITGLTVAIKMDIKKAMYAGVQVONQVKTICQLK 68
      118 LEDPEIGRPLGKGRGVNYLAREERKRIALLKTLFPGSOLKAEVGHQLRREVAIQLSHLR 177
Db      69 HPSILLEYNPFEDSNVYVLYLEKCHNGEMNRYLKORVYKPSSENAHFMQIITGMULYH 128
      178 HENILRLYGYHDSRYYLLIDYAPGELFRELQKCTR-FDDQSRNAYIKQALBALYCH 236
Qy      129 SHGILHDLTSLNLLTFNNMIKTLADFGLATQLMKPHKHYTLGGTPNYISPEIATRSAN 188
      237 SKKVIHRDIPKRNLLSGNSGLKIADFGMS--VHAPSSRRTLLCGTLDYLPPEMIEGRMH 294
Db      189 GLESDVWSLGCMPFTLLIGRPEDPTDVKTANKVNLADYEMPSFLSIEAKDLIHQILRR 248
      295 DEKVDLSLGLVLCYEFVVGKPPFETDTHQETVRIISKVEFOYPPYVSEBARDLVSKLKH 354
Qy      249 NPADRLSLSSVLDHPFMSRNS-STKSD 275
      355 NPNHRLPLKGVLEHPIWIKNSQLKKKD 382

```

```

RESULT 8
A47545
Protein kinase (EC 2.7.1.37) Plk - mouse
C/Species: Mus musculus (house mouse)
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C/Accession: A47545
R/Clay, P.J.; McEwen, S.J.; Bettoncello, I.; Wilks, A.F.; Dunn, A.R.
Proc. Natl. Acad. Sci. U.S.A. 90, 4882-4886, 1993

```

```

A/Title: Identification and cloning of a protein kinase-encoding mouse gene, Plk, related
A/Reference number: A47545; MVID:93281660; PMID:8099445
A/Accession: A47545
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-603 <CTA>
A/Cross-references: GB:L06144; NID:9309461; PIDN:AAA3948.1; PID:9309462
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C/Keywords: ATP; phosphotransferase
F:51-305/Domain: protein kinase homology <KIN>

```

```

Query Match      10.7%; Score 545; DB 2; Length 603;
Best Local Similarity 27.4%; Pred. No. 1.3e-14;
Matches 153; Conservative 93; Mismatches 192; Indels 120; Gaps 20;

Qy      15 GNILKGSPAGVYRAESITHTGLEVAIKMDIKKAMYAGVQVONQVKTICQLKPSILE 74
      56 GRPLGKGPFAKCFEISDADDTKEVFAGKIVPSLLKPKQKEMSMSEIHSLSAHQYVG 115
Db      75 LNYNPFEDSNVYVLYLEKCHNGEMNRYLKORVYKPSSENAHFMQIITGMULYHSHGILH 134
      116 FDDPFEDSDPFVYVLELCRRRSLEIHLKRR-KALTEPARAYLRQIVLGCQYLRNHYIH 174
Qy      135 RDLTSLNLLTFNNMIKTLADFGLATQLMKPHKHYTLGGTPNYISPEIATRSANGLSEDV 194
      175 RDLKGNLFLNEDLEVKIGDGLATKVEYEGERRKTLGGTPNYIAPVLSKSGHSFEVDV 234
Qy      195 WSLGCMFTLLIGRPEDPTDVKTANKVNLADYEMPSFLSIEAKDLIHQILRRNPADRL 254
      235 WSGICIMVTLVVGKPPFETSCIKETVYLIIKNGEVSIPGHNPVAAASLIQKMLQYDPTARP 294
Qy      255 SLSSVLDHPFMSRNSSTKSDLGTVEDSIDSGHATIS--FAIRASSTISGSLFD--K 309
      295 TIHEILNDEFTT-----SGVLPARLPITCLTTPREFSTAPSSLPSS 336
Qy      310 RRLI-----IGDPLNKMVTFPNKKSSTDPSSSGDG-----NSFYQWNGQETSN--SGR 357
      337 RKPLKLVNKGVENPLPDR-----PREKEPVYREINMAIECHLSLDLQQLTSVNASKSSER 392
Qy      358 GRVQDAEERH-----SRVLRAYSDBRSCTS-----NSQSQAITYMER 398
      393 GLVQOEAEEDACPIPIFWVSQKWD--YSDKYGLGYOLCNSVGVLPNDSTRLILY---- 445
Qy      399 CHSAEMLSVSRSGGSENERYSPTDNAN-----IFNPFKKTSSTSSGSPFRPNNG 451
      446 -NDDGSLQYLRDG-----TESYLTVSSHPSLSMKKTLILAYFR-----N 484
Qy      452 ALSNHL-C-PGKTPP-PADPTPQTEVQWF-----GNLQIN-----A 487
      485 YMSHLLKAGRNITPRRGDEIARLPYLRWFRTRSAIILHLSNGTVQVINFQDHTKILC 544
Qy      488 HLKRTYDSISPRDFO 505
      545 PLMAAVTY--INEXKDFQ 560

```

```

RESULT 9
A54596
Protein kinase - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C/Accession: A54596
R/Lake, R.J.; Jelinek, W.R.
Mol. Cell. Biol. 13, 7793-7801, 1993
A/Title: Cell cycle- and terminal differentiation-associated regulation of the mouse mRN
A/Reference number: A54596; MVID:94067140; PMID:7902533
A/Accession: A54596
A/Status: preliminary; translated from GB/EMBL/DBD/J
A/Molecule type: mRNA
A/Residues: 1-603 <RBS>
A/Cross-references: GB:L19558; NID:9403473; PIDN:AAA16071.1; PID:9403474
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
F:51-305/Domain: protein kinase homology <KIN>

```

Query Match	10.7%;	Score 545;	DB 2;	Length 603;
Best Local Similarity	27.2%;	Pred. No. 1.3e-14;		
Matches 152;	Conservative 95;	Mismatches 190;	Indels 122;	Gaps 20;

QY 15 GNLLGKSPFAGVYAAASIHITGLAEVAIKMIDKKAAYKAGMORVONEVYHQLKHPSTLE 74  
Db 56 GRPLGKGPAKCFEISIDADTKEVPAKGI VPSLSLLKHQKEMKMEISIHSLAHQHVIG 115  
QY 75 LYNFEDSNVYVLEVMCHNGEMNRYLKNRVKPSFENEARHPHQIITGMLYLHSHGLH 134  
Db 116 FHDFEEDSDFFVVLLECRRSRLLELHKRR-KALTEPEARYYLQIVLGCYLLHRNQYIH 174  
QY 135 RDLTSLMULLTRMNNIKIADFGLATOLKMPHEKRYTLCGTNNYISPELAITSAGLES DV 194  
Db 175 RDLQLGNLFMEDLEVXIGDFGLATKVEYEGERRKTYLGTENYLAPEVLSKKGHSFEVDV 234  
QY 195 WSLCCMFETYLIGRPEDTDTVKNTLNKVLAADYEMPSFLSEAKNDLIHOLLRRNPADRL 254  
Db 235 WSIICCIYTLVLGKPPETSCIKETYLRIKKNESIPRHINPVASLLOKMLQTDPRARP 294  
QY 255 SLSSVLDPFMSRNSSTKS KD LGTEVSDSGHATIS---TAITASSSTSI SGLSPD-K 309  
Db 295 TIHBLANDEFPT-----SGYIPARLPITCLTIPRFSTIAPSSLDPS 336  
QY 310 RRLI-----IGQPLFNKATVFPKOKSSTDPSSSGDC-----NSFTYQMGQETSN-SGR 357  
Db 337 RKPLKLVNKGVENLPDR---PREKEPVARENEALIECHLSDLLQOLTSVNASKPSER 392  
QY 358 GRVLODAEERPH-----SRYLRAVSSDRSGTS-----NSGQCATYTMER 398  
Db 393 GLVQGEAEADPACIPFMWSKMYD---YSDKYLGLCYQLCDNSVGLFNDSTRLLY---- 445  
QY 399 CHSAEMLSVSKRSGGGENERYSPTDNNAN-----IFNPKETKS---SSGSFERPD 448  
Db 446 -NDGDSLQYTERDG---TESYLVSSHPSHLSMKITLILNFRVYMSHLLKAGANYTPR 500  
QY 449 NNQALSNHLCPGKTPFPFADPTPOTETVQOWF-----GNLQIN----- 486  
Db 501 EGDELA-----RLPY-----LRTWFTRSAILIHLNSGTQVINEFODHTRLIL 543  
QY 487 AHLKTEYDISPNRDPQ 505  
Db 544 CPLMAAVTY--INEKRDPO 560

```

RESULT 10
G89287
protein H39E23.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: G89287
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:90565613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: G89287
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1246 c-to>
A:Cross-references: GB:chr_V; PIDN:CAE09532.1; PID:G3878100; GSPDB:GN00023; CESP:H39E23
C:Genetic:
A:Gene: H39E23.1
A:Map position: 5

```

Query Match	10.7%	Score 54.4.5	DB 2	Length 1246
Best Local Similarity	22.8%	Pred. No. 32-14		
Matches 252	Conservative 167	Mismatches 436	Indels 249	Gaps 38
12 PKVGNLLGKGSFAGYVRAESVHTGLEVAVIKMIDKKAMTKAGNQRVQNEVNIHQOLKSPS				
12				

Dh	60	YKLLKTIIGKGNFAVKLAKAVITGHEBAVAKIIDTALNPSSL-QGLPFEVXIMQDLDPN	118
Qy	72	ILELYNTFEDSNYYVLYLEMGCHGEANRYL-KRVYRPFSENEARHFMHOITGMLYHS	129
Dh	119	IVKLYQWMEIEQTLTYLYLEVASGEVEPYLVAHGMRK--EKEARAKPROVSAVOYLHS	175
Qy	130	HGILHRDLTSLNNLLTTRNMNKINDFGAQTQKMPHEKHYTLGCPNPTISPEI-ATRSAN	188
Dh	176	KNITHRDCKENLLDODMNKILDPFGHSNTSLGN-KLDFGCSPPYAAEBELBSGKYD	234
Qy	189	GLSEDSVSLGCMFTLLIGREPPTDVTYKNTANRYLVADYEMPSPSLIABQD.IHQLRR	248
Dh	235	GPEVDVMSLGYITLVIVSGSLPFGQGNLKEIREVLGKXNIPYMWSTDCENLLKKPLVI	294
Qy	249	NPADRLSLSSVLDHPFMSRN-----	268
Dh	295	NPQRSSLDINIMKORMMNVGEDEDLKFLTEPPPKQDIDECRIEKLQIFOLGFNKAILE	354
Qy	269	--SSTKCD-----LGYEDSLDSGHATISTALITASSTISIGSLPDKRRLIGOLP	319
Dh	355	SVEKEKEDIHATYLLIGERKSDMAEITMAQSLHSSTJNVSSL-----GQHPA	406
Qy	320	NKMT---VFPKNSKSTDFSSSGDGNSEFYQWGNQETSNSGGRVYODAEERPHSHYLRRA	376
Dh	407	CVITREHYVSSSAGSSASPSRYRSATATGASITGSA---LASANMAKH---QOS	459
Qy	377	YSSDRSGTNSQOAKTYTWERCHSAEMLSVKRSQGGENEER-----YS	421
Dh	460	SAAPSSSSSSRRSSQNDAAATAGCTGVVMSGTRHGGVQMBAOPTSRQATISLLPSPYK	519
Qy	422	PTDNNAIFN---FFKEKTSSSGSGSPR-----PDNNQALSNLCP--	459
Dh	520	PSMTWTOIAQIPLFNNRSTASTSAAPSTGISTGRIADPKGRIPLASTVAVQGRITATG	579
Qy	460	-----SKTPEPPADPPTQETTYQWQMGNIQINAHAKT-----TEYDISIPNDFOGHP	508
Dh	580	AVANNNGIIP-SHDDHAQOQOYNNQULSTSTMSGLKINTPAAGTAATASSSSSASTGA	638
Qy	509	DLQDQTSKNAWTDYVKVKNDSADSNANHSVKQOQNTMKYWTAL-HSKP-ETIQOECVFGS	564
Dh	639	PLQSGSGQISHAPEYVIREDDDEN-NSENQGNVPLIGVGQPOTSPAVQVPTEDATSS	697
Qy	565	DLPSQSKTRGMEP---PMQYQRTLRSTISPLVAHRLKIRQKTKKAAVSIIDSEVC	620
Dh	698	DKEQOQKASASETKESNPFIWQVLHNSLLSKSLDSSAATSYETPRPGIAGRRSESA	757
Qy	621	VELVKEASQOEYVKEVLOISD-----GNTI-----TIYFNGRGFPILADPPSPPT	667
Dh	758	ATPRRRQIQWVDARHLOTPDTRPRPHFEPTLLDROKRALYVSTASSRMTRGVLPTPTP	817
Qy	668	DNISRYSPDMLPEKYMRKYOYASHFVOLVBSKSPKIT-----YFTRYAK---711	
Dh	818	SNSTSSSFIYEPLRH-----VAASPDITTTPTKSTVTTSPYFRRTSPRM	864
Qy	712	-----CLIMENSPPADFEVWFDGVIKIKTEBFIQVIEKTSYTLKSSSEVNSLKEBELKM	767
Dh	865	VLIVLLLCDNG--LRIM--PSMIQSPSM-----PPSQMTAMESELK	903
Qy	768	YMDANEGHRICTALLESIISEBEKTRSAPEPFIIGRK-----PGST	810
Dh	904	SESQQTGGPPYATGPPORATSOQMSASATTNSANMGASGCGAAAAASATQOLGABEST	963
Qy	811	SS-----PKLASPPSPVDSNYPTRDRASFNRMYHSAASPTQAPILNPSWVTEGLGLT	864
Dh	964	GASQGYHPRKAPSSSSSSSTVPRHQHQLTNH-----ASFVTPSSYQPP--TSAVAVVT	1015
Qy	865	TTASGTDISSLKDCPKSAQOLKASFVVKVNGATQLTISAAVWQFPIDGSQLYVQA-CV	923
Dh	1016	STGMPTSSSSSA---PPRTRNRQOTRHGK-----TEKDQGG---DDSDDEIGETPGN	1061
Qy	924	SSISYTP---NGOTTRYGENEKL	944
Dh	1062	VSIQATGSPANNAEATITWSKSL	1085



```

A/Accession: T18611
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1192 <W11.1>
A/Cross-references: EMBL:281027; PIDN: CAB54179.1; GSPDB: GN00023; CESP: H39E23.1a
A/Experimental source: clone AH10
A/Accession: T18610
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-487,536-1192 <W11.2>
A/Cross-references: EMBL:281027; PIDN: CAB54178.1; GSPDB: GN00023; CESP: H39E23.1b
A/Experimental source: clone AH10
R/McMurray, A.
submitted to the EMBL Data Library, June 1997
A/Reference number: Z19696
A/Accession: T23144
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1192 <W11.3>
A/Cross-references: EMBL:296102; PIDN: CAB54263.1; GSPDB: GN00023; CESP: H39E23.1a
A/Experimental source: clone H39E23
A/Accession: T23143
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-487,536-1192 <W11.4>
A/Cross-references: EMBL:296102; PIDN: CAB54262.1; GSPDB: GN00023; CESP: H39E23.1b
A/Experimental source: clone H39E23
C/Genetics:
A/Genes: CESP: H39E23.1a; CESP: H39E23.1b
A/Map position: 5
A/Intons: 10/3; 33/3; 67/3; 139/2; 183/3; 451/3; 487/3; 535/3; 631/1; 825/2; 914/3; 992
C/Keywords: alternative splicing; ATP; phosphotransferase; serine/threonine-specific pro
F.1-1192/Product: probable serine/threonine-specific protein kinase, long splice form #8
F.1-487,536-1192/Product: probable serine/threonine-specific protein kinase, short splice
Query Match 10.4%; Score 527; DB 2; Length 1192;
Best Local Similarity 25.8%; Pred. No. 1,4e-13;
Matches 183; Conservative 108; Mismatches 252; Indels 166; Gaps 20;
QY 12 PTVGNLKGSGPAQVYRAESHTGLVAIKMDKAMTKAQVQVQVYKHCQLKHP 71
DB 170 YKLTKTIGGNFPAKVLAKHVIHGVHVAIKIIDKTLNPSSTL-QKLFREYKIMKQLDHP 228
QY 72 ILELVYFEDSNVYVLYLEMCHNGENRVL-KNRKPSSEARHFMQITGMVLYLS 129
DB 229 IYKLVQVMTBOTLYVLEVASGEVFDYLVHGRMK-EKARAKPRQIVSAVYLS 285
QY 130 HGIHLRDLTSLNLLTRNNNIKIADFGLATQQLKMPHEKHYTLGTPNYISPEI-ATRSAN 188
DB 286 KNIIRHDLKAEMLLDQDNKIKADFGFNTFSLGN-KLDTFGSPFYAAPFLFSKKYD 344
QY 189 GLESDVWSICMFEYTLIGRPPTDTVKNLKNVVLADYEMPSFLISEAKDLIHQLRR 248
DB 345 GEVVDVWSIGVLYLVVSGSLPFDQNLKELRERVLRGKRIIPFYMSTDCENLKKFLVY 404
QY 249 NPADRLSSVLDHPMSN----- 268
DB 405 NFORSSSLNIMKDRMMANQYEDDELKPIEPKQIDEGRIKLIQIFQLGNNKALILE 464
QY 269 --SSTKSD-----LGTVEDSIDSGHATITATITASSSTISGSLFDRRLILGQPLP 319
DB 465 SYEKEKFEIDIHATYLLIGERKSDMDASBITMAQSLSHSSINVSSL-----GQHPA 516
QY 320 NKMT--VPPKNSKSTDFSSGSDGNSFYQWGNQOETNSNGRGVYIDDAEERPHRYLRA 376
DB 517 GYITREHYVSSSSASGSSASPSRYSRSATATGASITAGSA---LASAANAQKH---QOS 569
QY 377 YSDRSGTNSQSOQATTYMERCHSABMLSVSKRSGGSEER-----YS 421
DB 570 SAAPSSGSSSSRSKSSONDAATAGGTVMVSGTRHGQVQWRAOPTSRQATISLQPPSYK 629
QY 422 PTDNNANIN-----PFKEKTSSSSGSFER-----PDNNQALSNHLCP-- 459

```

```

DB 630 PSSNTTQIAIPELPFNBNSTATSSAAQSTGILTRKRIADPKGRILPINSTAVQGHRTATG 669
QY 460 -----GKTPFPADPTQETETVQWFGNLQINAHLRKT-----TEYDISPRDQGH 508
DB 690 AYAANNNGIP-SHRDAQOOQVYNQLTSTMSKLINKPKAPAGTATATSSSSSATSTRA 748
QY 509 DLQKDTSK-NAWTDTRVKN-----SDASDNH 535
DB 749 PLQSGSQISHAPTEVIREDDENNENGNVPLIGVGPTSPAVQPTEDATSSSD 808
QY 536 SVKQNTKMYKTAHLSKPEITQEE-----CVFGSDPLSQOSTRG 575
DB 809 KEQOOQKASSSETPKESKPSMIMHSPMPSPQWMTAMESIKLSSESGQTGG 857

RESULT 14
S37928
probable purine nucleotide-binding protein YKL101w - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein YKL453
C/Species: Saccharomyces cerevisiae
C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 24-Sep-1999
C/Accession: S37928; S39084
R/Cheret, G.; Fukuhara, H.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Pallier, C.; Puzos
submitted to the Protein Sequence Database, March 1994
A/Reference number: S37920
A/Accession: S37928
A/Molecule type: DNA
A/Residues: 1-1518 <CHR>
A/Cross-references: EMBL:228101; NID: g486168; PIDN: CAB81941.1; PID: g486169; MIPS: YKL101w
A/Experimental source: strain S288C
R/Pallier, C.; Valens, M.; Puzos, V.; Fukuhara, H.; Cheret, G.; Sor, F.; Bolotin-Fukuhara
Yeast 9, 1149-1155, 1993
A/Title: DNA sequence analysis of a 17 kb fragment of yeast chromosome XI physically loc
protein kinase.
A/Reference number: S39084; MUID: 94078677; PMID: 8256524
A/Accession: S39084
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-1518 <PAL>
A/Cross-references: EMBL: X71133; NID: g431205; PIDN: CAB50456.1; PID: g431215
A/Experimental source: strain S288C
C/Genetics:
A/Genes: SGD: HSL1
A/Cross-references: SGD: S0001584; MIPS: YKL101w
A/Map position: 11L
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C/Keywords: ATP; P-loop; purine nucleotide binding; serine/threonine-specific protein kin
F.79-369/Domain: protein kinase homology <KIN>
F.79-86/Region: nucleotide-binding motif A (P-loop)
F.87-95/Region: protein kinase ATP-binding motif
F.85/Binding site: ATP/GTP (Yrs) #status predicted
Query Match 10.4%; Score 526.5; DB 2; Length 1518;
Best Local Similarity 22.3%; Pred. No. 1.9e-13;
Matches 253; Conservative 182; Mismatches 429; Indels 271; Gaps 42;
QY 7 EKIEDPKGNLKGSGPAQVYRAESHTGLVAIKMDKKA-----MYKAG 53
DB 76 DTVGPMKDKTYLKGSGGRVLRKAKMETQALAIKIVPKKAFVHCSSNGTVPNSYSSM 135
QY 54 VQR-----VQNEVKIKCHQLKHPSSILELVYFEDSNVYVLYLEM 91
DB 136 VTSNVSSPSIARSHNSHOTPYGERIRIVMKLISHNNVALFPWENKSLYLVLVE 195
QY 92 CHNGEMRYLKORVXPFSENEAHFMHQITGMVLYHSHGILHRDITLSNLLI-TRNNMI 150
DB 196 VDCGEFLDYLVSKGK-LPEREAIHYKQIVEGVYCHSFNICHRLDKPENLLLDKKQNRRI 254
QY 151 KTADFGLATQQLKMPHEKHYTLGTPNYISPEIAT-PSAAGLSDVWSLGMFPYTLILGRP 209
DB 255 KIADFGMAA-LEIPNLKLTQSGSPHYASPELVMGVPHGGSDVWSCGIVLFAALTGH 313
QY 210 PFDITVKNLKNVVLADYEMPSFLISEAKDLIHQLRRNPADRLSSVLDHP----- 264

```

```

Db      314 PFNDNMTKLLKLVGSGKQMPSPNLSSEARDLISKILVIDPERKITTOELIKPLIKKD 373
Qy      265 -----MSKNSSTKSD-----LCVDSISDGAH--TISPAITASST---SIS 303
Db      374 DLPEVKVLRMRKDNNAFGKNSDHLHLNNVSPSIVTLHSKGRIDESIRSLQILMHGVS 433
Qy      304 GSLPDKRLLIGQPLPNKMTVPF-----KVKSTDPSSSGDGSFYTQNGNOSTNSGRG 358
Db      434 RELITAK--LQKPMEEKLFYLLLOKORHSISLSSSEKKAT-----478
Qy      359 RVIQDAERPHSRYLRAVSDRSSTNSQAKTYTMRCHSAEMLSVSKSG--GENE 417
Db      479 -----ESSVNERLEYSK--TANNTGLRSENNVKTILHSLEHSDSTVQNNALITGVNT 533
Qy      418 ERYSPTDNNANINFNFKETSSSSSGFEPDPNNQALSNHL--CGKTPPFPAD-----468
Db      534 EIMAPV-----LAQSQFISINTLSQPSDKAAEAVTLPALPIFNASSSRIFRNS 584
Qy      469 ----PYQTEVQQWFGNLIQINAHLRKTEYDSSISPNRDFOGHPDLQKDTSKNAWTDKV 524
Db      585 YNISRSRSLRLSLRSRLSLSTSLSTRETYHDNEMF-----LPQLPKSPKSYLSRAKI 638
Qy      525 KKNSDASNAHSVQKQNTKMTALHSKPEIILQECVFGSDPLSEQSKTRGKMPWGYON 584
Db      639 HASPSTKSIHKSLSRKNIAATVAARITLQNSAKRSLSLQISIKRS-----LNL 688
Qy      585 RFLRSTISPLVHARKLPIROKTKKANVSLDS-----EVCVELV-----624
Db      689 NDILVFPDLPKSK--KPASENVNKSSEPHLESSEDFELLCDQILFGNALDILLEEEDNE 746
Qy      625 KEVASQEVYKAVLQISDQNTI---TIYYPNGRGFPPLA-----DRPSPDNTISRYAF 675
Db      747 KERDTPQROQNDTKSSADFTTISGVSTKNENGPEPTIKKQNFMSYKSENNSGLS 806
Qy      676 DULPEYTKRKQYASFPVOLVNSKSPKITTYFTRYAKCIIMENSFGADFEVWFYDGYKHK 735
Db      807 PFIFEK---ENTLSSSYLEOKPKRALSDINTSFNKNKQO-----EGRIERK 851
Qy      736 TDFDITQVLEKTKSYTLK---SESEVNSLKEE---IKMYMD--HANEGRICLAEISLIS 787
Db      853 KIQROLOLQKNDPRLKPIQHOELRVNLSLPMDOGPISLDPRRNVIQSPVNSKVESLQ 911
Qy      788 EEEKTRSAPEPPIIIG-----RKPGSTSPKALSPPSVDSNY--PTRDRASPNR 836
Db      912 GLKFKKEPASHWTHRGSLFMSEHVEDEKPVKASD-----VSLESSVPLTIVATSR 964
Qy      837 WMMHGAASPT-QAPF-----LNPMTYNEGILGTTASGTD-----871
Db      965 DPSVLAESSTIQKPMLSLPSFPLNTSM--TFKNLSQITLADGDKHLSPVQNSRVSAMSH 1023
Qy      872 -ISSNSLKDCLPKSAQLKSFVK-----NVGMATQL-----TSGAV 907
Db      1024 PLRKQSAKISLTPRNLNANLSVKKNQSGPSYLSNDLDGISDMTFAMEIPNTYTAQAI 1083
Qy      908 WYQFND-----GSQLVVQAGVSSISYTS--PNCQTTYGENEKLKDY 947
Db      1084 QLMNNDTNNKINTSPKASSFTKEKVIKSAAYISKKEKPDNDTNY-----IPDY 1133

```

## RESULT 15

serine/threonine-specific protein kinase p101 (BC 2.7.1.-) - fission yeast (Schizosaccha  
 C/Species: Schizosaccharomyces pombe  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jul-2000  
 C/Accession: J38254; T45128  
 R/Brown, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
 submitted to the EMBL Data Library, August 1995  
 A/Reference number: Z21781  
 A/Accession: J38254  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-683 <BRO>

A/Cross-references: EMBL:Z98559; PIDN:CA811167.1; GSPDB:GN00066; SPDB:SPAC23C11.16  
 A/Experimental source: strain 972h; cosmid C23C11  
 R/Ohkura, H.; Hagan, I.M.; Glover, D.M.  
 Genes Dev. 9, 1059-1073, 1995  
 A/Title: The conserved Schizosaccharomyces pombe kinase p101, required to form a bipolar  
 A/Reference number: Z22921; MUID:95562899; PMID:7744248  
 A/Accession: T45128  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-683 <OHK>  
 A/Cross-references: EMBL:X85758; MID:9887640; PIDN:CA59766.1; PID:9887641  
 A/Experimental source: strain 972 derivative  
 C/Genetics:  
 A/Map position: 1  
 C/function:  
 A/Description: required to form a bipolar spindle and early in the regulatory cascade the  
 A/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
 C/Keywords: phosphotransferase; protein kinase

Query Match 10.1%; Score 514.5; DB 2; Length 683;  
 Best Local Similarity 24.4%; Pred. No. 2.4e-13;  
 Matches 176; Conservative 106; Mismatches 247; Indels 191; Gaps 22;

```

Qy      18 LKGSFAGVYRAESIHTGLJEVALIKMIDKKAMVQVONENVKIHCOLKHPSTILELYN 77
Db      47 IGGGFACCFVKD--NYGNITAAKVIKRSLONDKTKLKGELIKHOSMHPNIVGID 105
Qy      78 YFEDSNVYVLEMCNCHGNNRNYLKNRVKPPSENEARHPHQIITGMULYHSHGLHRDL 137
Db      106 CFEDSTNIYLLILELCEHKSIMELLKR--KQLEPEVRVIMQILGALKYMKKKVYHRDL 164
Qy      138 TLSNILLTRNNIKIADRGATOLKMPHEKHYTLGTPNYISPIRA--TRAHGLESVW 195
Db      165 KLGNTLDESNNVAKIGDFGLALLMDDEERMTTICGPNYIAPILFENSKGSHFEVDLV 224
Qy      196 SLGCMFYTLIGRPFPDVTYKNTLNKVLADYEMPS--FLSIKADLIHOLLRRNPADR 253
Db      225 SAGVVMYVALLIGKPPFOQKVKTYRKIKANSYFPSPVVDLSARAKOLISLTHHPDIR 284
Qy      254 LSLSSVLDPHPPMSRNSSTKSKDOLGTVEDSIDGHATTSTAITASSSTISGSLPDKRLL 313
Db      285 PSIDIDVDFE-----FHTGYMAST-----304
Qy      314 IGQPLPMQWTFPKPKKST-----DSSSGDGSFYTQNG-----NOETSN---354
Db      305 LPDELHSMPIWPSQSCKSFQRLNDFVASGVGFNSAGVEKOKPYALTDEVDNDRI 364
Qy      355 -----SGRGVIO---DAERPHSRYLRAVSDRSSTNS-----QSOA---391
Db      365 LPSVLSPRDRVNPWKIGPFTKVPKSLSTLHAARKSTDSLSRVKVLREESQFVPT 424
Qy      392 KTYMERCHSAEMLSVSKRSGGGENERYSPTDNNANINFNFKKTKSSSGSFERRPDNQ 451
Db      425 KSAVTEQVEPQLI---RSLSANTVSRLSKVGMKSDIWSVYKTLKIG-----471
Qy      452 ALSNHLCPGKTPFPADPTQTEVQWFGNLIQINAHLRKTEYDSS-----499
Db      472 -----MALEATHALTSEDADSEPVLPITTKVNDY 500
Qy      500 PNRDFOGHPDLQKDTSKNAWTDTKYKKNSDASNAHSVQKQNT-----MKYMTALH 550
Db      501 SNKYGGLGVLSDSEVGHPNDTSLPSADBEVEYVALHPKDTIKPIYIPASVPSIR 560
Qy      551 SKPITIQECVFGSDPLSEQSKTRGMPWGYONRTLRSLTSPVLRKLPFRKT-----606
Db      561 SKIQLHFKFSYMQNSKAVQDESFEKP--KNST--SYTMFMQHYLR--TRQALMFR 614
Qy      607 -----KAAVS-----LIDSEVCEVLKVEYASQ-----EVKVEVLO 638
Db      615 SNGIFQNPFLDHRKAVVISTARKIIVLDKREKREVELPQZASAFSEDLRSRKTIYRETL 674

```

Mon Sep 27 12:38:15 2004

Search completed: September 23, 2004, 21:02:30  
Job time : 27 secs

us-10-026-021-2.rpr

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 20:59:21 : Search time 22 Seconds

(without alignments)  
2276.236 Million cell updates/sec

Title: US-10-026-021-2

Perfect score: 5078  
Sequence: 1 MATCGEKEDPKVGNLJLK.....KQGLSSILMFSPNPNH 970

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgnt2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
2: /cgnt2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgnt2\_6/prodata/2/1aa/6A\_COMB.pep:\*  
4: /cgnt2\_6/prodata/2/1aa/6B\_COMB.pep:\*  
5: /cgnt2\_6/prodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgnt2\_6/prodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3927.5	77.3	925	1 US-08-252-995D-4	Sequence 4, Appl1
2	3927.5	77.3	925	2 US-08-834-108-4	Sequence 4, Appl1
3	1926	37.9	464	1 US-08-252-995D-6	Sequence 6, Appl1
4	1926	37.9	464	2 US-08-834-108-6	Sequence 6, Appl1
5	1883.5	37.1	416	1 US-08-252-995D-2	Sequence 2, Appl1
6	1883.5	37.1	416	2 US-08-834-108-2	Sequence 2, Appl1
7	1370	27.0	273	1 US-08-252-995D-10	Sequence 10, Appl1
8	1370	27.0	273	2 US-08-834-108-10	Sequence 10, Appl1
9	590.5	11.6	607	2 US-08-878-989-15	Sequence 15, Appl1
10	590.5	11.6	607	3 US-09-272-796-15	Sequence 15, Appl1
11	578.5	11.4	271	1 US-08-252-995D-11	Sequence 11, Appl1
12	578.5	11.4	271	2 US-08-834-108-11	Sequence 11, Appl1
13	561.5	11.1	272	1 US-08-252-995D-12	Sequence 12, Appl1
14	561.5	11.1	272	2 US-08-834-108-12	Sequence 12, Appl1
15	560.5	11.0	685	2 US-08-878-989-1	Sequence 1, Appl1
16	560.5	11.0	685	3 US-09-136-282-2	Sequence 2, Appl1
17	560.5	11.0	685	3 US-09-272-796-1	Sequence 1, Appl1
18	545	11.0	685	3 US-09-505-744-1	Sequence 1, Appl1
19	545	10.7	603	4 US-09-111-311C-26	Sequence 26, Appl1
20	533	10.5	603	4 US-09-198-122-2	Sequence 2, Appl1
21	525.5	10.3	272	1 US-08-252-995D-14	Sequence 14, Appl1
22	525.5	10.3	272	2 US-08-834-108-14	Sequence 14, Appl1
23	508.5	10.0	403	2 US-08-755-728-4	Sequence 4, Appl1
24	508.5	10.0	403	2 US-08-974-655-4	Sequence 4, Appl1
25	508.5	10.0	403	3 US-09-283-011-4	Sequence 4, Appl1
26	498	9.8	722	4 US-08-817-8328-32	Sequence 32, Appl1
27	495.5	9.8	722	4 US-09-984-890-4	Sequence 4, Appl1

28	490	9.6	275	1 US-08-252-995D-13	Sequence 13, Appl1
29	490	9.6	275	2 US-08-834-108-13	Sequence 13, Appl1
30	487	9.6	724	4 US-09-984-890-2	Sequence 2, Appl1
31	485.5	9.6	344	2 US-08-755-728-3	Sequence 3, Appl1
32	485.5	9.6	344	2 US-08-974-655-3	Sequence 3, Appl1
33	485.5	9.6	344	3 US-09-283-011-3	Sequence 3, Appl1
34	480	9.5	347	2 US-09-016-000-1	Sequence 1, Appl1
35	479.5	9.4	745	4 US-09-523-849-36	Sequence 36, Appl1
36	464	9.1	729	2 US-08-677-298-2	Sequence 2, Appl1
37	464	9.1	729	4 US-09-523-849-33	Sequence 33, Appl1
38	462.5	9.1	633	3 US-08-557-006C-43	Sequence 43, Appl1
39	454.5	9.0	556	2 US-09-016-000-4	Sequence 4, Appl1
40	454.5	9.0	556	4 US-09-156-793D-2	Sequence 2, Appl1
41	450	8.9	1037	4 US-09-428-711A-21	Sequence 21, Appl1
42	449	8.8	793	4 US-09-523-849-32	Sequence 32, Appl1
43	446.5	8.8	556	4 US-09-800-960-4	Sequence 4, Appl1
44	446.5	8.8	556	4 US-10-096-960-4	Sequence 4, Appl1
45	445.5	8.8	1203	4 US-09-799-875-5	Sequence 5, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-252-995D-4  
Sequence 4, Application US/08252995D  
Patent No. 6565050A  
GENERAL INFORMATION:  
APPLICANT: Dennie, James W  
APPLICANT: Heffernan, Mike  
APPLICANT: Fode, Carol  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERSKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3J2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,995D  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdyak, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-96  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 925 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-252-995D-4  
Query Match 77.3%; Score 3927.5; DB 1; Length 925;  
Best Local Similarity 78.6%; Pred. No. 8.4e-296;  
Matches 763; Conservative 76; Mismatches 83; Indels 49; Gaps 9;  
QY 1 MATCGEKEDPKVGNLJLKSGSFAGYRABSIHTGLEVAIKMIDKAMTKAGVQVQV 60  
DB 1 MAACIGIEDPFGVGNLKGSPAGYRABSIHTGLEVAIKMIDKAMTKAGVQVQV 60  
QY 61 VKIHQGLKPSIELYVNFEDSNVYVLVLEMGCHGEMNRYLKNRVKPFSEARHFMHQI 120

```

Db      61 VKIHQKHPVLELYNFEDNNVYVLEWCHGENRKYLNKMKCFSEBEAHFMHQI 120
Qy      121 ITGMLYHSHGILHRDLTSLNLLTRMNNIKIADFGATOLKMPHEKHYTLGTPNYISP 180
Db      121 ITGMLYHSHGILHRDLTSLNLLTRMNNIKIADFGATOLKMPHEKHYTLGTPNYISP 180
Qy      181 EIATRSAGLESADVSLGCMFYTLIGRPEDTDVKNLKVVLADYEMPSFLISAKD 240
Db      181 EIATRSAGLESADVSLGCMFYTLIGRPEDTDVKNLKVVLADYEMPSFLISAKD 240
Qy      241 LIHOLLERNPADRLSLSSVLDHPMSRNSSTKSKDGLTVEDSIDSGHATTSTAITSST 300
Db      241 LIHOLLERNPADRLSLSSVLDHPMSRNSSTKSKDGLTVEDSIDSGHATTSTAITSST 300
Qy      301 SLSGLSLD-RLLVGGPLPKKITVFOKNKNSSDP--SSGDGNSNFTQMGNEBQENRGRG 358
Db      301 SLSGLSLD-RLLVGGPLPKKITVFOKNKNSSDP--SSGDGNSNFTQMGNEBQENRGRG 358
Qy      359 RVIODAEERPHSRYLRAAYSSDSGTNSQSOAKTYTMEKCHSAEMLSVSRSGGGENEE 418
Db      359 RVIODAEERPHSRYLRAAYSSDSGTNSQSOAKTYTMEKCHSAEMLSVSRSGGGENEE 418
Qy      419 RYSPDNNANIFNPFKKTSSSSGSRPRPNNOALSNHLCPGTPPPADPTOTETVQO 478
Db      411 -----LDENQHSNHLCLGKTPPPADPTOTETVQO 442
Qy      479 WFGNLOINAHLRKTEYDSISPNRDFQGHPLQKOTSKNAMTDTKYKNSDASDNASHYK 538
Db      443 WFGNLOINAHLRKTEYDSISPNRDFQGHPLQKOTSKNAMTDTKYKNSDASDNASHYK 501
Qy      539 QQNTMKTMTLHAKPELIIQOECVFGSDPLSEQSKTRGMEPPMGYQNRTLRSITSPVAHR 598
Db      502 QLSAMKTYMSAHNKKPEVMPQEP--GLHPHSEQSKNSMESTLQYQKPTLRISITSPVAHR 559
Qy      599 LKPIROKTKKAVYSIIDSEVCYELVEXASOEYVEVLOISDGMITITTYNNGRGRP 658
Db      560 LKPIROKTKKAVYSIIDSEVCYELVEXASOEYVEVLOISDGMITITTYNNGRGRP 619
Qy      659 LADRPSPDNTISRSYFDMLEPKYMKRYOYASRFVLRVSKSPKITTYFTRAKCLIMENS 718
Db      620 LADRPSPDNTISRSYFDMLEPKYMKRYOYASRFVLRVSKSPKITTYFTRAKCLIMENS 679
Qy      719 PGADFEVWYDYGIKIKHTDFOVIEKTKGKSTYLKSESEVNIKEETIKMYMDHANGRI 778
Db      680 PGADFEVWYDYGIKIKHTDFOVIEKTKGKSTYLKSESEVNIKEETIKMYMDHANGRI 739
Qy      779 CLALBSIISEBEKRTSAPFFPIIIGRKPGSTSPKALSPPSVDSNYPTRDRASPNRV 838
Db      740 CLALBSIISEBEKRTSAPFFPIIIGRKPGSTSPKALSPPSVDSNYPTRDRASPNRV 798
Qy      839 MMSAASPLOAPILNPMVTNEGILTTASGTDISSNLKDCLPKSAQLLKSVFYANVGM 898
Db      799 MMSAASPLOAPILNPMVTNEGILTTASGTDISSNLKDCLPKSAQLLKSVFYANVGM 853
Qy      899 ATQOLTSAGVWVOPNDOSQLVQAGVSSISYTSPNGQTRRYGENEKLPDVTKLOCLSSI 958
Db      854 ATQOLTSAGVWVOPNDOSQLVQAGVSSISYTSPNGQTRRYGENEKLPDVTKLOCLSSI 913
Qy      959 LLMFSNPTNPF 969
Db      914 LLMFSNPTNPF 924

```

RESULT 2  
US-08-834-108-4  
; Sequence 4, Application US/08834108  
; Patent No. 5976893

GENERAL INFORMATION:  
; APPLICANT: Dennis, James W  
; APPLICANT: Heffernan, Mike  
; APPLICANT: Fode, Carol  
; TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE

```

; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERSKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,108
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurdydk, Linda M
; REGISTRATION NUMBER: 34,971
; REFERENCE/DOCKET NUMBER: 3153-210
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 925 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-834-108-4

```

Query Match 77.3%; Score 3927.5; DB 2; Length 925;  
Best Local Similarity 78.6%; Pred. No. 8.4e-296;

Matches 763; Conservative 76; Mismatches 83; Indels 49; Gaps 9;

```

Qy      1 NATIGEKIEDPFKYNLLGKSPAGVYRAESIHGLEVALIMIDKKAMKYAMVQVONE 60
Db      1 MACIGERIEDPFKYNLLGKSPAGVYRAESIHGLEVALIMIDKKAMKYAMVQVONE 60
Qy      61 VKIHQKHPVLELYNFEDNNVYVLEWCHGENRKYLNKMKCFSEBEAHFMHQI 120
Db      61 VKIHQKHPVLELYNFEDNNVYVLEWCHGENRKYLNKMKCFSEBEAHFMHQI 120
Qy      121 ITGMLYHSHGILHRDLTSLNLLTRMNNIKIADFGATOLKMPHEKHYTLGTPNYISP 180
Db      121 ITGMLYHSHGILHRDLTSLNLLTRMNNIKIADFGATOLKMPHEKHYTLGTPNYISP 180
Qy      181 EIATRSAGLESADVSLGCMFYTLIGRPEDTDVKNLKVVLADYEMPSFLISAKD 240
Db      181 EIATRSAGLESADVSLGCMFYTLIGRPEDTDVKNLKVVLADYEMPSFLISAKD 240
Qy      241 LIHOLLERNPADRLSLSSVLDHPMSRNSSTKSKDGLTVEDSIDSGHATTSTAITSST 300
Db      241 LIHOLLERNPADRLSLSSVLDHPMSRNSSTKSKDGLTVEDSIDSGHATTSTAITSST 300
Qy      301 SLSGLSLD-RLLVGGPLPKKITVFOKNKNSSDP--SSGDGNSNFTQMGNEBQENRGRG 358
Db      301 SLSGLSLD-RLLVGGPLPKKITVFOKNKNSSDP--SSGDGNSNFTQMGNEBQENRGRG 358
Qy      359 RVIODAEERPHSRYLRAAYSSDSGTNSQSOAKTYTMEKCHSAEMLSVSRSGGGENEE 418
Db      359 RVIODAEERPHSRYLRAAYSSDSGTNSQSOAKTYTMEKCHSAEMLSVSRSGGGENEE 418
Qy      419 RYSPDNNANIFNPFKKTSSSSGSRPRPNNOALSNHLCPGTPPPADPTOTETVQO 478
Db      411 -----LDENQHSNHLCLGKTPPPADPTOTETVQO 442
Qy      479 WFGNLOINAHLRKTEYDSISPNRDFQGHPLQKOTSKNAMTDTKYKNSDASDNASHYK 538
Db      443 WFGNLOINAHLRKTEYDSISPNRDFQGHPLQKOTSKNAMTDTKYKNSDASDNASHYK 501
Qy      539 QQNTMKTMTLHAKPELIIQOECVFGSDPLSEQSKTRGMEPPMGYQNRTLRSITSPVAHR 598

```

Db	502	QUSAMKTVYSAHHHAKKEVWFQEP--GLHPHSBQSKHSMESTLQYQKPTLRKSTSPILAAH	559
Qy	599	LKPIRQKTKKAAVVSILDSBECVELVKERASOBVYKVELQISSDGNITITYYPNNGRGP	658
Db	560	LKPIRQKTKKAAVVSILDSBECVELLRBCASGVYKVELQISSDGNITITYYPNNGRGP	619
Qy	659	LADRPSPSTDNISRSFNDLPEKTYWRKYQYASBFVQVLSKSPKITYPTRYAKCIMENS	718
Db	620	LADRPPLPTDNISRSFNDLPEKYWRKYQYASFIQVLSKPKITYPTRYAKCIMENS	679
Qy	719	PGADEWVFYQGVKTHKTEDEFLQVIRKTKGSKYTLKSESVNSLKEEIKYMDHANEGHRI	778
Db	680	PGADEWVFYQGAHKTHKTEENLHILEKTIYSTLNKNEVENTSLKEEIKYMDHANEGHRI	739
Qy	779	CLALESIISSEERKTRSAFPFPIIIGRKGPSTSSPALSPSPVDSNYPTRDRASFNRW	838
Db	740	CLISESVISSEERKRGRGSFFPIIYORKKGNTPSPALAPF--VDSGCCGBOASASRLS	798
Qy	839	MHSAASPQAPILNDSMYTNEGILGTTTASGTDISSNSLKDCLPKSAQILLKSVFVXVNGW	898
Db	799	VNSAFAFPQSPCLSPSTVYVEGLGHATATXNGVSS-----LPKSAQILLKSVFVXVNGW	853
Qy	899	ATQLTSGAVWQVQFNDSQQLVWQAGVSSISYTSBNGQTTTRYGENEKKIPDYIKQQLCLLSI	958
Db	854	ATQLTSGAVWQVQFNDSQQLVWQAGVSSISYTSBQGTTRYGENEKKIPDYIKQQLCLLSI	913
Qy	959	LLMPSNPPTNPF	969
Db	914	LLMPSNPPTNPF	924

```

1      RESULT 3
2      US-08-252-995D-6
3      Sequence 6, Application US/08252995D
4      Patent No. 5650501
5      GENERAL INFORMATION:
6      APPLICANT: Dennie, James W
7      APPLICANT: Heffernan, Mike
8      APPLICANT: Fode, Carol
9      TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASEE
10     NUMBER OF SEQUENCES: 14
11     CORRESPONDENCE ADDRESS:
12     ADDRESSEE: BERSKIN & PARR
13     STREET: 40 King Street West
14     City: Toronto
15     STATE: Ontario
16     COUNTRY: Canada
17     ZIP: M5H 3Y2
18
19     COMPUTER READABLE FORM:
20     MEDIUM TYPE: Floppy disk
21     COMPUTER: IBM PC compatible
22     OPERATING SYSTEM: PC-DOS/MS-DOS
23     SOFTWARE: Patentin Release #1.0, Version #1.30
24
25     CURRENT APPLICATION DATA:
26     APPLICATION NUMBER: US/08/252,995D
27     FILING DATE: 02-JUN-1994
28     CLASSIFICATION: 536
29     ATTORNEY/AGENT INFORMATION:
30     NAME: Kurdydyk, Linda M
31     REGISTRATION NUMBER: 34,971
32     REFERENCE/DOCKET NUMBER: 3153-96
33     TELECOMMUNICATION INFORMATION:
34     TELEPHONE: (416) 364-7311
35     TELEFAX: (416) 361-1398
36
37     INFORMATION FOR SEQ ID NO: 6:
38     SEQUENCE CHARACTERISTICS:
39     LENGTH: 464 amino acids
40     TYPE: amino acid
41     TOPOLOGY: linear
42
43     MOLECULE TYPE: protein
44
45     US-08-252-995D-6

```

Query Match	37.9%	Score 1926	DB 1	Length 464
Best Local Similarity	84.9%	Pred. No. 5.1e-141		
Matches 372	Conservative 30	Mismatches 30	Indels 6	Gaps 5

Qy	1	MATIGEXIEDEPKGNLLGKSPFAGVPAEASHTGLJVAIKMIDKKAMKAGMVRQNE	60
Dp	1	MAATIGERIEDEPKGNLLGKSPFAGVPAEASHTGLJVAIKMIDKKAMKAGMVRQNE	60
Qy	61	VKIHQOLKHPISLELYNFYEDSNYYLVLEMCHEGEMNRYYLKNRYKPSSENAHFHMOI	120
Dp	61	VKIHQOLKHPISLELYNFYEDSNYYLVLEMCHEGEMNRYYLKNRKKPFSEARAHFHMOI	120
Qy	121	ITGMVLYHSHGIIHRDLTLSNILLTRNNNNIKIADGGLATOLKMPHEKHTLCTPRYISR	180
Dp	121	ITGMVLYHSHGIIHRDLTLSNILLTRNNNNIKIADGGLATOLKMPHEKHTLCTPRYISR	180
Qy	181	EIATRSAGHGESDYSWLSGCMFYTLIGRPEDTDVTKNTLNVVLADYEMPSFLISBAQD	240
Dp	181	EIATRSAGHGESDYSWLSGCMFYTLIGRPEDTDVTKNTLNVVLADYEMPAFLISBAQD	240
Qy	241	LIHOLLRRNPADRLSLSSVLDHPFMSRKSSTYKSXOLGTYVEDSIDBGHATISATITASSST	300
Dp	241	LIHOLLRRNPADRLSLSSVLDHPFMSRNPSPKSKXGVGYVEDSMDGHAITSTITASSST	300
Qy	301	SISGSLFPIKRLLLIGQPLFNQWTFYPPKKSSTDFSSSDGNSPFTQWGN--OETSNSGRG	358
Dp	301	SISGSLLD--KULLVGQPLPKKITTFQKKNSSDF--SSGDGNSFCYQWGNPBEAANSRGG	358
Qy	359	RVIODAERPHRSRYLRRAYSSDRSGTSSQOAKTYTHERCHSAEMLSVSKSGGGENEE	418
Dp	359	RVIDEAERPHRSRYLRRAYSSDRASPNS--QSRAKTYSVBRCHSVEMLSKPRRS--LDENOH	416
Qy	419	RYSPTDNNANLNFNFKEK	436
Dp	417	RYSPTKSNVNLTSILNTK	434

RESULT 4  
 US-08-834-108-6  
 Sequence 6, Application US/08834108  
 Patent No. 5976893  
 GENERAL INFORMATION:  
 APPLICANT: Dennis, James W  
 APPLICANT: Heffernan, Mike  
 APPLICANT: Fode, Carol  
 TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BERSKIN & PARR  
 STREET: 40 King Street West  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5H 3Y2  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/834,108  
 FILING DATE:  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kurdydyk, Linda M  
 REGISTRATION NUMBER: 34,971  
 REFERENCE/DOCKET NUMBER: 3153-210  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 364-7311  
 TELEFAX: (416) 361-1398  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 464 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-834-108-6

Query Match 37.9%; Score 1926; DB 2; Length 464;  
Best Local Similarity 84.9%; Pred. No. 5,1e-141;  
Matches 372; Conservative 30; Mismatches 30; Indels 6; Gaps 5;

QY 1 MATCIGKIDPKVGNLKGSPAGVYRAESIHGTEVAIKMIDKKAMTKAGVQVONE 60  
DB 1 MAACIGRIEDFKVGNLKGSPAGVYRAESIHGTEVAIKMIDKKAMTKAGVQVONE 60  
QY 61 VKIHQOLKHPSTLELYNYPEDSNVYVLEMCNGENNRYLKVRKPFSEARHPMHQI 120  
DB 61 VKIHQOLKHPSTLELYNYPEDSNVYVLEMCNGENNRYLKVRKPFSEARHPMHQI 120  
QY 121 ITGMLYLSHGILHRDLTSLNLLTRNNMIKIADFGLATQLMNPKHEKHYTLGCTPNYISP 180  
DB 121 ITGMLYLSHGILHRDLTSLNLLTRNNMIKIADFGLATQLMNPKHEKHYTLGCTPNYISP 180  
QY 181 EATRSAGLESVDWSLSCMFTYLLIGRPPTDTVKNLTKVLADEMPFLSTEAKD 240  
DB 181 EATRSAGLESVDWSLSCMFTYLLIGRPPTDTVKNLTKVLADEMPFLSTEAKD 240  
QY 241 LIHQOLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATISATIASST 300  
DB 241 LIHQOLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATISATIASST 300  
QY 301 SISGSLFDRKRLILGQPLPKMTVPKNSSTDPSSSGDGNSTYTWGN--QETSNSGRG 358  
DB 301 SISGSLD-RLVLGQPLPKMTVPKNSSTDPSSSGDGNSTYTWGN--QETSNSGRG 358  
QY 359 RVIQDAERPHSRYLRRAYSSDRSGTSNSQOATYTMERCHSABMLSVSKSGGGENE 418  
DB 359 RVIQDAERPHSRYLRRAYSSDRSGTSNSQOATYTMERCHSABMLSVSKSGGGENE 418  
QY 419 RVSPTNNANINFPFEK 436  
DB 419 RVSPTKSNVNLVLSLNTK 434

## RESULT 5

US-08-252-995D-2  
Sequence 2, Application US/08252995D  
Patent No. 5650501  
GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Heffernan, Mike  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERSKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,995D  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydyk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-96  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 416 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-252-995D-2

Query Match 37.1%; Score 1883.5; DB 1; Length 416;  
Best Local Similarity 87.7%; Pred. No. 8,6e-138;  
Matches 362; Conservative 27; Mismatches 19; Indels 5; Gaps 4;

QY 1 MATCIGKIDPKVGNLKGSPAGVYRAESIHGTEVAIKMIDKKAMTKAGVQVONE 60  
DB 1 MAACIGRIEDFKVGNLKGSPAGVYRAESIHGTEVAIKMIDKKAMTKAGVQVONE 60  
QY 61 VKIHQOLKHPSTLELYNYPEDSNVYVLEMCNGENNRYLKVRKPFSEARHPMHQI 120  
DB 61 VKIHQOLKHPSTLELYNYPEDSNVYVLEMCNGENNRYLKVRKPFSEARHPMHQI 120  
QY 121 ITGMLYLSHGILHRDLTSLNLLTRNNMIKIADFGLATQLMNPKHEKHYTLGCTPNYISP 180  
DB 121 ITGMLYLSHGILHRDLTSLNLLTRNNMIKIADFGLATQLMNPKHEKHYTLGCTPNYISP 180  
QY 181 EATRSAGLESVDWSLSCMFTYLLIGRPPTDTVKNLTKVLADEMPFLSTEAKD 240  
DB 181 EATRSAGLESVDWSLSCMFTYLLIGRPPTDTVKNLTKVLADEMPFLSTEAKD 240  
QY 241 LIHQOLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATISATIASST 300  
DB 241 LIHQOLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATISATIASST 300  
QY 301 SISGSLFDRKRLILGQPLPKMTVPKNSSTDPSSSGDGNSTYTWGN--QETSNSGRG 358  
DB 301 SISGSLD-RLVLGQPLPKMTVPKNSSTDPSSSGDGNSTYTWGN--QETSNSGRG 358  
QY 359 RVIQDAERPHSRYLRRAYSSDRSGTSNSQOATYTMERCHSABMLSVSKSGGGENE 411  
DB 359 RVIQDAERPHSRYLRRAYSSDRSGTSNSQOATYTMERCHSABMLSVSKSGGGENE 410

## RESULT 6

US-08-834-108-2  
Sequence 2, Application US/08834108  
Patent No. 5976893  
GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Heffernan, Mike  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERSKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/834,108  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydyk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-210

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 416 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-834-108-2

Query Match 37.1%; Score 1883.5; DB 2; Length 416;  
Best Local Similarity 87.7%; Pred. No. 8.6e-138;  
Matches 362; Conservative 27; Mismatches 19; Indels 5; Gaps 4;

QY 1 MATCGEKEDPKVGNLGGSPAGYVRAESITGLEVAKIMIDKAMKAGVQVONE 60  
DB 1 MACIGERIEDPKVGNLGGSPAGYVRAESITGLEVAKIMIDKAMKAGVQVONE 60  
QY 61 VTIHQQLKPSVLELYNYFEDNNYVYLVEMCHNGEMNRYLKVRVPESENEARHFMHQI 120  
DB 61 VTIHQQLKPSVLELYNYFEDNNYVYLVEMCHNGEMNRYLKVRVPESENEARHFMHQI 120  
QY 121 ITGMLYHSHGILHRDLTSLNLLTRNNMIKIADFGIATQLKMPHEKHVTLGCTPNYIS 180  
DB 121 ITGMLYHSHGILHRDLTSLNLLTRNNMIKIADFGIATQLKMPHEKHVTLGCTPNYIS 180  
QY 181 EIATSSAGLESVDWMLGCMFTYLLIGRPFDYVNTLNKVVADYEMPSFLSEAD 240  
DB 181 EIATSSAGLESVDWMLGCMFTYLLIGRPFDYVNTLNKVVADYEMPSFLSEAD 240  
QY 241 LHHLLRRPADRLSSVLDHPFMSRNSSTKSDIGTEDSIDGHATTSTITASSST 300  
DB 241 LHHLLRRPADRLSSVLDHPFMSRNSSTKSDIGTEDSIDGHATTSTITASSST 300  
QY 301 SISGSLDFDKRRLILGQPLPNKMTVPFNKNSSTDPSSSGDGNSFTYTMGN--QETSNGRG 358  
DB 301 SISGSLDFDKRRLILGQPLPNKMTVPFNKNSSTDPSSSGDGNSFTYTMGN--QETSNGRG 358  
QY 359 RVIQDAEERPHSRYLRRAYSSDRSGTSSGQSQAKYTTMERCHSAEMLSKRS 411  
DB 359 RVIQDAEERPHSRYLRRAYSSDRSGTSSGQSQAKYTTMERCHSAEMLSKRS 411

RESULT 7  
US-08-252-995D-10  
Sequence 10, Application US/08252995D  
Patent No. 5650501  
GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Heffernan, Mike  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERSKIN & PAR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydk, Linda M  
REGISTRATION NUMBER: 34,971

REFERENCE/DOCKET NUMBER: 3153-96  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 273 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-252-995D-10

Query Match 27.0%; Score 1370; DB 1; Length 273;  
Best Local Similarity 94.5%; Pred. No. 3e-98;  
Matches 258; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 5 IGEKIDPKVGNLGGSPAGYVRAESITGLEVAKIMIDKAMKAGVQVONEVKIH 64  
DB 1 IGERIEDPKVGNLGGSPAGYVRAESITGLEVAKIMIDKAMKAGVQVONEVKIH 60  
QY 65 COLKHPSTIELYNYFEDSNVYVLEMCNGEMNRYLKVRVPESENEARHFMHQIITGM 124  
DB 65 COLKHPSTIELYNYFEDSNVYVLEMCNGEMNRYLKVRVPESENEARHFMHQIITGM 120  
QY 125 LTIHSHGILHRDLTSLNLLTRNNMIKIADFGIATQLKMPHEKHVTLGCTPNYISPEIAT 184  
DB 125 LTIHSHGILHRDLTSLNLLTRNNMIKIADFGIATQLKMPHEKHVTLGCTPNYISPEIAT 180  
QY 185 RSHAGLESVDWMLGCMFTYLLIGRPFDYVNTLNKVVADYEMPSFLSEADLIHQ 244  
DB 185 RSHAGLESVDWMLGCMFTYLLIGRPFDYVNTLNKVVADYEMPSFLSEADLIHQ 240  
QY 245 LRRNPADRLSSVLDHPFMSRNSSTKSDIG 277  
DB 245 LRRNPADRLSSVLDHPFMSRNSSTKSDIG 273

RESULT 8  
US-08-834-108-10  
Sequence 10, Application US/08834108  
Patent No. 5976893  
GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Heffernan, Mike  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERSKIN & PAR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-210  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398



```

/ APPLICATION NUMBER: US/09/272,796
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/878,989
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J J
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0321 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 15:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 607 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 1827450
/ US-09-272-796-15

Query Match 11.6%; Score 590.5; DB 3; Length 607;
Best Local Similarity 37.5%; Pred. No. 3,1e-37;
Matches 115; Conservative 61; Mismatches 106; Indels 25; Gaps 4;

QY 15 GNILGKSPAGVYRAESHTGLEVALIKMIDKAMVYAGVQVQVONVVKIHCOLKPSILE 74
DB 26 GLLLGKSPAGVYRAESHTGLEVALIKMIDKAMVYAGVQVQVONVVKIHCOLKPSILE 85
QY 75 LNVYFEDSNVYVIVEMCHNGENNRKLYKRVKPFSENEARHFMHOITIGMLYHSHGILH 134
DB 86 FSHHEDADNIVYIVELCKRSKSLAHMKAR-HTLLEPEVRYVYRQLLSGKYLHOGILH 144
QY 135 RDLTSLNLLITRNMIKIDFGLATOLKMPHEKHYTLGTPNYISPEIATRSAGHESDV 194
DB 145 RDLKGNFETTERMEKVDPLGLAALBEPBQKKTICGTPNYVAEVLRLQHGFEADV 204
QY 195 WSLGCMFYTLIGRPEDTDVKNLTKVYLADYEMPSFLSIKADLIHQLLRNPADRL 254
DB 205 WSLGCMFYTLIGRPEDTDVKNLTKVYLADYEMPSFLSIKADLIHQLLRNPADRL 264
QY 255 SLSSVLDHPFMRNSSTKSDGTVDSDSGHATISTAITASSST-----SIGG 304
DB 265 SLSDIIRHDFPK-----GYTPDRL-----PISCVTVDLTPNPASRLFAKVTK 310
QY 305 SLFDKRR 311
DB 311 SLFGKKK 317

RESULT 11
US-08-252-995D-11
/ Sequence 11, Application US/08252995D
/ Patent No. 5650501
/ GENERAL INFORMATION:
/ APPLICANT: Dennis, James W
/ APPLICANT: Heffernan, Mike
/ APPLICANT: Fode, Carol
/ TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BERSKIN & PARR
/ STREET: 40 King Street West
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5H 3Y2
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk

```

```

/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/252,995D
/ FILING DATE: 02-JUN-1994
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kurydyk, Linda M
/ REGISTRATION NUMBER: 34,971
/ REFERENCE/DOCKET NUMBER: 3153-96
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (416) 364-7311
/ TELEFAX: (416) 361-1398
/ INFORMATION FOR SEQ ID NO: 11:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 271 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: Drosophila melanogaster
/ US-08-252-995D-11

Query Match 11.4%; Score 578.5; DB 1; Length 271;
Best Local Similarity 42.5%; Pred. No. 7,4e-37;
Matches 108; Conservative 53; Mismatches 92; Indels 1; Gaps 1;

QY 12 FVGNILGKSPAGVYRAESHTGLEVALIKMIDKAMVYAGVQVQVONVVKIHCOLKPS 71
DB 7 YKMRPFKSPAGVYRAESHTGLEVALIKMIDKAMVYAGVQVQVONVVKIHCOLKPS 66
QY 72 ILELVYFEDSNVYVIVEMCHNGENNRKLYKRVKPFSENEARHFMHOITIGMLYHSHG 131
DB 67 IYKFNHYFEDSONIVYIVELCKRSKSMELAKRR-KSITEFECRYVYQIIGVKYLHNR 125
QY 132 IIRDLTSLNLLITRNMIKIDFGLATOLKMPHEKHYTLGTPNYISPEIATRSAGHES 191
DB 126 IIRDLKGNLFLNDLHAKIDFGLATRIEYGERKKTLCGTANYIAEILTKGHSTP 185
QY 192 SDVWSLSCMFYTLIGRPEDTDVKNLTKVYLADYEMPSFLSIKADLIHQLLRNPA 251
DB 186 VDIWSLSCMFYTLIGRPEDTDVKNLTKVYLADYEMPSFLSIKADLIHQLLRNPA 245
QY 252 DRSLSSVLDHPFM 265
DB 246 SRPAIGQLNFEFL 259

RESULT 12
US-08-834-108-11
/ Sequence 11, Application US/08834108
/ Patent No. 5976893
/ GENERAL INFORMATION:
/ APPLICANT: Dennis, James W
/ APPLICANT: Heffernan, Mike
/ APPLICANT: Fode, Carol
/ TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
/ NUMBER OF SEQUENCES: 14
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BERSKIN & PARR
/ STREET: 40 King Street West
/ CITY: Toronto
/ STATE: Ontario
/ COUNTRY: Canada
/ ZIP: M5H 3Y2
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/834,108  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-210  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 271 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Drosophila melanogaster  
US-08-834-108-11

Query Match 11.4%; Score 578.5; DB 2; Length 271;  
Best Local Similarity 42.5%; Pred. No. 7,4e-37;  
Matches 108; Conservative 53; Mismatches 92; Indels 1; Gaps 1;

QY 12 PKVGNLKGSGFAGVYRAESHTGLEVAIKMTDKKAMYKAGVORVONEVKIHCOLKHPSS 71  
DB 7 YKRMRFPGGGRKACEIITIDVEDVDPAKIVSKMLIKHNOKEKRAQETTHRSLSNHR 66  
QY 72 ILELVYFEDSNVYVLEMCCHNGENRRLKRVKPFSENEARHFMHQIITGMVLYSHSG 131  
DB 67 IYVFNYPEDSQIYIVLELCKRSMWELHKRR-KSITEFECRYIYQIIGVXYLHNR 125  
QY 132 ITHRDLTSLNLTTRMNTKIADFGLATQKMPHEKHTLCTPNVTSPELTARSAHGE 191  
DB 126 ITHRDLTSLNLTTRMNTKIADFGLATQKMPHEKHTLCTPNVTSPELTARSAHGE 185  
QY 192 SDVWSLGCMPYTLILGRPFEDTDTVNTLNKVLADYEMPSFLSEAKDLIHLRRNPA 251  
DB 186 VDIWSIGCMYTLIVGQPFETKTLADTYSKIKCKEYRVSYLKRRAADVMVIMLQNPPE 245  
QY 252 DRLSSVLDHPM 265  
DB 246 SRPAIGQLNFERL 259

RESULT 13  
US-08-252-995D-12  
Sequence 12, Application US/08252995D  
Patent No. 5650501  
GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Hefernan, Mike  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERESKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,995D  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:

NAME: Kurdydk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-96  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 272 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
US-08-252-995D-12

Query Match 11.1%; Score 561.5; DB 1; Length 272;  
Best Local Similarity 41.6%; Pred. No. 1,6e-35;  
Matches 104; Conservative 46; Mismatches 99; Indels 1; Gaps 1;

QY 15 GNULGKGSFAGVYRAESHTGLEVAIKMTDKKAMYKAGVORVONEVKIHCOLKHPSS 74  
DB 11 GAVLKGSGFAGVYRAESHTGLEVAIKMTDKKAMYKAGVORVONEVKIHCOLKHPSS 70  
QY 75 LNVYFEDSNVYVLEMCCHNGENRRLKRVKPFSENEARHFMHQIITGMVLYSHSG 134  
DB 71 FYHVPEDKNITVLELFCGRSMHILKAR-KVLEPEVRYVLRQVSGIKTLHEGEIIL 129  
QY 135 RDLTSLNLTTRMNTKIADFGLATQKMPHEKHTLCTPNVTSPELTARSAHGE 194  
DB 130 RDLKGNFPIINAMELVKGDVGLAARLEPLERRRTICTPNVTSPELTARSAHGE 189  
QY 195 WSLGCMYTLILGRPFEDTDTVNTLNKVLADYEMPSFLSEAKDLIHLRRNPA 254  
DB 190 WSLGCMYTLILGRPFEDTDTVNTLNKVLADYEMPSFLSEAKDLIHLRRNPA 249  
QY 255 SLSSVLDHPF 264  
DB 250 SLDDIIRHDF 259

RESULT 14  
US-08-834-108-12  
Sequence 12, Application US/08834108  
Patent No. 5976893  
GENERAL INFORMATION:  
APPLICANT: Dennis, James W  
APPLICANT: Hefernan, Mike  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERESKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/834,108  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydk, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-210  
TELECOMMUNICATION INFORMATION:



;  
;  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:

Job time : 24 secs

Search completed: September 23, 2004, 21:02:56  
Job time : 24 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model1

Run on: September 24, 2004, 14:49:50 ; Search time 198 Seconds  
(without alignments)  
8164.505 Million cell updates/sec

Title: US-10-026-021-1

Perfect score: 2913  
Sequence: 1 atggcgacctgcgcgcggga.....cgacctatattcatga 2913

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/prodata/2/ina/5A\_COMB.seq: \*  
4: /cgn2\_6/prodata/2/ina/5B\_COMB.seq: \*  
5: /cgn2\_6/prodata/2/ina/5A\_COMB.seq: \*  
6: /cgn2\_6/prodata/2/ina/5B\_COMB.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2911.4	99.9	3937	4 US-09-620-312D-280	Sequence 280, Appl
2	1879	64.5	3447	1 US-08-252-995D-3	Sequence 3, Appl
3	1879	64.5	3447	2 US-08-834-108-3	Sequence 3, Appl
4	970.4	33.3	1600	1 US-08-252-995D-5	Sequence 5, Appl
5	970.4	33.3	1600	2 US-08-834-108-5	Sequence 5, Appl
6	924.2	31.7	1453	1 US-08-252-995D-1	Sequence 1, Appl
7	924.2	31.7	1453	2 US-08-834-108-1	Sequence 1, Appl
8	154.2	5.3	381	4 US-09-016-434-649	Sequence 649, Appl
9	134.4	4.6	2770	2 US-08-878-989-8	Sequence 8, Appl
10	134.4	4.6	2770	3 US-09-272-796-8	Sequence 8, Appl
11	134.4	4.6	2789	3 US-09-136-282-3	Sequence 3, Appl
12	134.4	4.6	2789	3 US-09-505-744-3	Sequence 3, Appl
13	132.4	4.5	2783	3 US-09-136-282-1	Sequence 1, Appl
14	132.4	4.5	2783	3 US-09-505-744-1	Sequence 1, Appl
15	131.4	4.5	2169	4 US-09-016-434-1147	Sequence 1147, Appl
16	118	4.1	2033	4 US-09-660-925B-10	Sequence 10, Appl
17	113.2	3.9	2198	2 US-08-755-728-2	Sequence 2, Appl
18	113.2	3.9	2198	2 US-08-974-655-2	Sequence 2, Appl
19	113.2	3.9	2198	2 US-09-283-011-2	Sequence 2, Appl
20	113.2	3.9	2253	4 US-09-660-925B-3	Sequence 3, Appl
21	109.4	3.8	1244	2 US-09-198-122-1	Sequence 1, Appl
22	104	3.6	1244	2 US-08-755-728-1	Sequence 1, Appl
23	104	3.6	1244	2 US-08-974-655-1	Sequence 1, Appl
24	104	3.6	1244	2 US-09-283-011-1	Sequence 1, Appl
25	103.6	3.6	1244	2 US-09-359-161-4	Sequence 4, Appl
26	99	3.4	1891	3 US-09-289-466-1	Sequence 1, Appl
27	97.4	3.3	1864	4 US-09-156-793D-1	Sequence 1, Appl

28	97.4	3.3	1890	3 US-09-289-466-2	Sequence 2, Appl
29	97.4	3.3	1929	2 US-09-016-000-10	Sequence 10, Appl
30	96	3.3	1281	2 US-09-016-000-7	Sequence 7, Appl
31	89.8	3.1	1590	4 US-09-554-726A-4	Sequence 4, Appl
32	89.8	3.1	1732	4 US-09-430-564-1	Sequence 1, Appl
33	89.8	3.1	2257	4 US-09-554-726A-6	Sequence 6, Appl
34	89.2	3.1	2556	4 US-09-817-310-1	Sequence 1, Appl
35	87.8	3.0	2754	3 US-09-429-322-3	Sequence 3, Appl
36	87	3.0	1637	2 US-08-966-316-10	Sequence 10, Appl
37	86	3.0	4438	4 US-09-566-921-81	Sequence 81, Appl
38	82.4	2.8	2698	2 US-08-677-298-1	Sequence 1, Appl
39	82.2	2.8	2076	4 US-09-554-726A-11	Sequence 11, Appl
40	82.2	2.8	2827	4 US-09-554-726A-11	Sequence 11, Appl
41	82.2	2.8	2827	4 US-09-554-726A-20	Sequence 20, Appl
42	80.6	2.8	2298	4 US-09-975-326-3	Sequence 3, Appl
43	80.6	2.8	2298	4 US-10-217-357-3	Sequence 3, Appl
44	80.6	2.8	2301	4 US-09-975-326-1	Sequence 1, Appl
45	80.6	2.8	2301	4 US-10-217-357-1	Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-09-620-312D-280  
; Sequence 280, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonphong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yungting  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinshast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 280  
; LENGTH: 3937  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (334)..(3246)  
; US-09-620-312D-280

Query Match 99.9%; Score 2911.4; DB 4; Length 3937;  
Best local similarity 100.0%; Pred. No. 0;  
Matches 2912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGACCTGCATCGCGGAGAAATCGAGATTGGAATCTGCTGTGTA 60  
DB 334 ATGGCGACCTGCATCGCGGAGAAATCGAGATTGGAATCTGCTGTGTA 393

QY 61 GAATCATTTGCTGGTGTCTACAGAGCTAGTCCATTCACTGGTTTGAAGTTGCAATC 120  
Db 394 GGAATCATTTGCTGGTGTCTACAGAGCTAGTCCATTCACTGGTTTGAAGTTGCAATC 453  
QY 121 AAAATGATAGTAAAGAAAGCCATGTACAAAGCAGAAATGGTACAGAGTCCAAATGAG 180  
Db 454 AAAATGATAGTAAAGAAAGCCATGTACAAAGCAGAAATGGTACAGAGTCCAAATGAG 513  
QY 181 GTGAAAAATACATTTGCCAATTGAAACATCTTCTATCTGTGGAGCTTTATTAATCTATTTTGA 240  
Db 514 GTGAAAAATACATTTGCCAATTGAAACATCTTCTATCTGTGGAGCTTTATTAATCTATTTTGA 573  
QY 241 GATAGCAATATATGTATCTGTATTAAGAAATGTGCCATATAGAGAAATGAACAGGTAT 300  
Db 574 GATAGCAATATATGTATCTGTATTAAGAAATGTGCCATATAGAGAAATGAACAGGTAT 633  
QY 301 CTAAAGAAATAGATGAAACCCCTTCTCAGAAATAAGAGCTGCACTTCATGACCAAGATC 360  
Db 634 CTAAAGAAATAGATGAAACCCCTTCTCAGAAATAAGAGCTGCACTTCATGACCAAGATC 693  
QY 361 ATCAGAGGATGTGTATCTTCAATCTCATGTATATCTACACGGGAGCTCAACATTTCT 420  
Db 694 ATCAGAGGATGTGTATCTTCAATCTCATGTATATCTACACGGGAGCTCAACATTTCT 753  
QY 421 AACCTCTTACTGACTCGTAATATGAACATCAAGATTGCTGATTTTGGCTGGCACTCAA 480  
Db 754 AACCTCTTACTGACTCGTAATATGAACATCAAGATTGCTGATTTTGGCTGGCACTCAA 813  
QY 481 CTGAAAAATGCACATGAAAAAGCATATATCTATATGTGGAACTCTTAACTTACATTTTCA 540  
Db 814 CTGAAAAATGCACATGAAAAAGCATATATGTGGAACTCTTAACTTACATTTTCA 873  
QY 541 GAAATTTGCCACTCGAAGTGCACATGGCTTGAATCTGATGTGTGCTCCCTGGGCTGTATG 600  
Db 874 GAAATTTGCCACTCGAAGTGCACATGGCTTGAATCTGATGTGTGCTCCCTGGGCTGTATG 933  
QY 601 TTTTATACATTAATCTTATGTGGAGAACCACTTGAACATCTGACATCAAGTCAAGAACATTA 660  
Db 934 TTTTATACATTAATCTTATGTGGAGAACCACTTGAACATCTGACATCAAGTCAAGAACATTA 993  
QY 661 AATAAAGTATATGTCAGATTAATGAATGCAATCTTTTGTGCAATTAAGGCAAGAAC 720  
Db 994 AATAAAGTATATGTCAGATTAATGAATGCAATCTTTTGTGCAATTAAGGCAAGAAC 1053  
QY 721 CTATATCACCAAGTACTTGTGTGAATAATCCAGCAGATCGTTTAACTGTCTTCAATATG 780  
Db 1054 CTATATCACCAAGTACTTGTGTGAATAATCCAGCAGATCGTTTAACTGTCTTCAATATG 1113  
QY 781 GACCAATCCTTTATATGCCCCGGAATTTCTTCAACAAAAAGTAAAGATTAGAACTGTGAA 840  
Db 1114 GACCAATCCTTTATATGCCCCGGAATTTCTTCAACAAAAAGTAAAGATTAGAACTGTGAA 1173  
QY 841 GACTCAATGATATGAGGAGATGCCAATTTCTACTGCAATTAAGCTCTTCCAGTACC 900  
Db 1174 GACTCAATGATATGAGGAGATGCCAATTTCTACTGCAATTAAGCTCTTCCAGTACC 1233  
QY 901 AGTATATGTTGATAGTTATTTTGAACAAAGAGACTTTGATTTGTCAAGCACTCCCAAT 960  
Db 1234 AGTATATGTTGATAGTTATTTTGAACAAAGAGACTTTGATTTGTCAAGCACTCCCAAT 1293  
QY 961 AAAATGACTGTATTTCCAAAGAAATAAAGTTGCACTGATTTTCTTCTTCAAGAGATGA 1020  
Db 1294 AAAATGACTGTATTTCCAAAGAAATAAAGTTGCACTGATTTTCTTCTTCAAGAGATGA 1353  
QY 1021 AACAGTTTATATCTCAAGTGGGAAATCAAGAAACAGATATATGGAAGGGAAGAGTA 1080  
Db 1354 AACAGTTTATATCTCAAGTGGGAAATCAAGAAACAGATATATGGAAGGGAAGAGTA 1413  
QY 1081 ATTCAAGATGCAAGAAAGGCCAATTTCTGATACCTTCTGTAAGCTTATTTCTCTGAT 1140  
Db 1414 ATTCAAGATGCAAGAAAGGCCAATTTCTGATACCTTCTGTAAGCTTATTTCTCTGAT 1473  
QY 1141 AGATCTGGCACTTTAATATAGTCAATCTCAAGCAAAAAATATATGATGAACAGATGTCA 1200

Db 1474 AGATCTGGCACTTTAATATAGTCTCAAGCAAAAAATATATGATGAACAGATGTCA 1533  
QY 1201 TCAGCAAAATGCTTCAAGTGTCCAAAAGATCAGAGAGAGTGAATAAGAGAGTAC 1260  
Db 1534 TCAGCAAAATGCTTCAAGTGTCCAAAAGATCAGAGAGAGTGAATAAGAGAGTAC 1593  
QY 1261 TCAGCCAGAGACAACTAGGCCCAATTTTAACTTTTAAAGAAAGATCCAGTAC 1320  
Db 1594 TCAGCCAGAGACAACTAGGCCCAATTTTAACTTTTAAAGAAAGATCCAGTAC 1653  
QY 1321 TCTGATCTTTTGAAGAAGCTGATTAACAATCAAGCACTCCCAATCATCTTTTGCAGGA 1380  
Db 1654 TCTGATCTTTTGAAGAAGCTGATTAACAATCAAGCACTCCCAATCATCTTTTGCAGGA 1713  
QY 1381 AAAAATCTTTTCCATTTGCAAGCCGACACTCGACTGAAACCGTACAAAGTGT 1440  
Db 1714 AAAAATCTTTTCCATTTGCAAGCCGACACTCGACTGAAACCGTACAAAGTGT 1773  
QY 1441 GGGAAATCTGCAAAATTAATGCTCATTTTAAGAAAACTAGTAATAGACATCAAGCCCA 1500  
Db 1774 GGGAAATCTGCAAAATTAATGCTCATTTTAAGAAAACTAGTAATAGACATCAAGCCCA 1833  
QY 1501 AACCGGACCTTCCAGGGCCATCCAGATTTGCAAGAGACACATCAAAAAATGCTGAGT 1560  
Db 1834 AACCGGACCTTCCAGGGCCATCCAGATTTGCAAGAGACACATCAAAAAATGCTGAGT 1893  
QY 1561 GATTAACAAAGTCAAAAAAGACTGTAGTCTTGATTAATGCACTTGTGTAACAGCA 1620  
Db 1894 GATTAACAAAGTCAAAAAAGACTGTAGTCTTGATTAATGCACTTGTGTAACAGCA 1953  
QY 1621 AATAACATGAATAATATATGATGCACTTCAAGTAAACCTGAGATPATCAACAAAGAT 1680  
Db 1954 AATAACATGAATAATATATGATGCACTTCAAGTAAACCTGAGATPATCAACAAAGAT 2013  
QY 1681 GTTTTGGCTCAGATCTCTTTTGTGAACAGAGCAAGACTAGGGGTATGAGCCACATGG 1740  
Db 2014 GTTTTGGCTCAGATCTCTTTTGTGAACAGAGCAAGACTAGGGGTATGAGCCACATGG 2073  
QY 1741 GGTATATCAATTCGTAACATTAAGAAAGCTTCACTCCGTGTGTGCTCAAGSTTAAA 1800  
Db 2074 GGTATATCAATTCGTAACATTAAGAAAGCTTCACTCCGTGTGTGCTCAAGSTTAAA 2133  
QY 1801 CCAATCAGACAGAAAAACCAAAAAAGCTGTGTGAGACATACTTGAATTCAGAGAGTGT 1860  
Db 2134 CCAATCAGACAGAAAAACCAAAAAAGCTGTGTGAGACATACTTGAATTCAGAGAGTGT 2193  
QY 1861 GTGAGCTTGTAAAGAGATGCAATCTCAGAAATATGTGAAAGAGTTCTTCAATATCT 1920  
Db 2194 GTGAGCTTGTAAAGAGATGCAATCTCAGAAATATGTGAAAGAGTTCTTCAATATCT 2253  
QY 1921 AGTATGGAATATAGATCACTATTTATTAATCAATGTGTGTAGAGTTTCCCTTGTCT 1980  
Db 2254 AGTATGGAATATAGATCACTATTTATTAATCAATGTGTGTAGAGTTTCCCTTGTCT 2313  
QY 1981 GATAGACCACTCAGCTACTGACCAACATCAGTATGATACGCTTTGCAATTTTACAGAA 2040  
Db 2314 GATAGACCACTCAGCTACTGACCAACATCAGTATGATACGCTTTGCAATTTTACAGAA 2373  
QY 2041 AAATATCTGGGAAATATCAATATGCTTCCAGGTTTGTACAGCTTGTATGAATCTAAATCT 2100  
Db 2374 AAATATCTGGGAAATATCAATATGCTTCCAGGTTTGTACAGCTTGTATGAATCTAAATCT 2433  
QY 2101 CCCAAATCACTTATTTTAAAGATATGCTAAATGCAATTTTGAATGAGAAATCCCTGCT 2160  
Db 2434 CCCAAATCACTTATTTTAAAGATATGCTAAATGCAATTTTGAATGAGAAATCCCTGCT 2493  
QY 2161 GCTGATTTTGAAGTTTGTATATGATGGGTAAATAACAACAAGAGATTTCAAT 2220  
Db 2494 GCTGATTTTGAAGTTTGTATATGATGGGTAAATAACAACAAGAGATTTCAAT 2553  
QY 2221 CAGGTGATTTGAAAAAGAGAGGAGTCTTACATTTTAAAGTGAAGTGAATATAGC 2280

Db 2554 CAGGTGATGTAAGAGACAGGAGAGCTTACCTTTAAGGAGTAAGTAAAGC 2613  
Qy 2281 TTGAAAAGAGAGTAAATATGTATATGACATGCTTAATAGAGGTCAATGTTGTTA 2240  
Db 2614 TTGAAAAGAGAGTAAATATGTATATGACATGCTTAATAGAGGTCAATGTTGTTA 2673  
Qy 2341 GCATCGAATCCATATTTTCAAGAGAGGAAAGAACTAGAGAGTCCCTTTTCCCA 2400  
Db 2674 GCATCGAATCCATATTTTCAAGAGAGGAAAGAACTAGAGAGTCCCTTTTCCCA 2733  
Qy 2401 ATATCATAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460  
Db 2734 ATATCATAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2793  
Qy 2461 TCTGTGATTTCAATTTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520  
Db 2794 TCTGTGATTTCAATTTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2853  
Qy 2521 AGTGTGCTTTCTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
Db 2854 AGTGTGCTTTCTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2913  
Qy 2581 CTGTGCTTACACATCAAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Db 2914 CTGTGCTTACACATCAAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2973  
Qy 2641 CTGTGCTTACACATCAAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
Db 2974 CTGTGCTTACACATCAAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3033  
Qy 2701 CAGTTACTAGTGGAG 2760  
Db 3034 CAGTTACTAGTGGAG 3093  
Qy 2761 GCAGAGAGTCTTCTATCAAGTTATACCTCAAGAGAGAGAGAGAGAGAGAGAGAG 2820  
Db 3094 GCAGAGAGTCTTCTATCAAGTTATACCTCAAGAGAGAGAGAGAGAGAGAGAGAG 3153  
Qy 2821 AATGAAAATTTACAG 2880  
Db 3154 AATGAAAATTTACAG 3213  
Qy 2881 ATGTTTCTATCGAG 2913  
Db 3214 ATGTTTCTATCGAG 3246

RESULT 2  
US-08-252-995D-3  
Sequence 3, Application US/08252995D  
Patent No. 5650501  
GENERAL INFORMATION:  
APPLICANT: Dennie, James W  
APPLICANT: Heffernan, Mike  
APPLICANT: Fode, Carol  
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERESKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,995D  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:  
NAME: Kurdyak, Linda M  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 3153-96  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3447 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE: Mus musculus  
ORGANISM: Mus musculus  
DEVELOPMENTAL STAGE: Lymphoid cDNA library  
IMMEDIATE SOURCE:  
LIBRARY: Murine Lymphoid  
CLONE: MGA-resistant chop clones  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..205  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 206..2980  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 2981..3447  
US-08-252-995D-3

Query Match 64.5%; Score 1879; DB 1; Length 3447;  
Best Local Similarity 80.9%; Pred. No. 0; Mismatches 410; Indels 147; Gaps 9;  
Matches 2362; Conservative 0;

Qy 1 ATGGCAGCTGATCGGGAG 60  
Db 206 ATGGCAGCTGATCGGGAG 265  
Qy 61 GGATCAATTTGCTGTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
Db 266 GGATCAATTTGCTGTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 325  
Qy 121 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
Db 326 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 385  
Qy 181 GTGAATATCATTTGCAATTTGAAACATCTTCTATCTTGAAGCTTATATATTTTGA 240  
Db 386 GTGAATATCATTTGCAATTTGAAACATCTTCTATCTTGAAGCTTATATATTTTGA 445  
Qy 241 GATGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
Db 446 GATGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 505  
Qy 301 CTAAAGATAGAGTGAACCTTTCTAGAAATGAAGCTTCGACACTTCAATGACCGATC 360  
Db 506 CTAAAGATAGAGTGAACCTTTCTAGAAATGAAGCTTCGACACTTCAATGACCGATC 565  
Qy 361 ATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
Db 566 ATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 625  
Qy 421 AACCTCTACGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
Db 626 AACCTCTACGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 685  
Qy 481 CTGAAGATGCAATGAAAG 540  
Db 686 CTGAAGATGCAATGAAAG 745  
Qy 541 GAAATGCACTGAG 600

Db 746 GAAATTCGAACCTGAGTGCACATGCACTTGATATTTTGTCAATTTGGGCTGTAG 805  
Qy 601 TTTTATACATTACTTATTCGGAGACCACTTTCGACACTGACACAGTCAAGAACATTA 660  
Db 806 TCTTATACCTTACTTATTCGAGACCACTTTTGCACATGACACAGTCAAGAACATTA 865  
Qy 661 AATTAAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAG 720  
Db 866 AACAAAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAG 925  
Qy 721 GTTATTCACACGATTCCTGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 780  
Db 926 GTTATTCACACGATTCCTGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 985  
Qy 781 GACCAATCTTTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 840  
Db 986 GACCAATCTTTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1045  
Qy 841 GACTCAATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 900  
Db 1046 GACTCAATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1105  
Qy 901 AGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 960  
Db 1106 AGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1162  
Qy 961 AAAATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1020  
Db 1163 AAAATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1219  
Qy 1021 AACAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1074  
Db 1220 AGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1279  
Qy 1075 AAGATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1134  
Db 1280 AAGATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1339  
Qy 1135 TCGATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1194  
Db 1340 TCGATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1396  
Qy 1195 TGTCACTGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1254  
Db 1397 TGTCACTGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1434  
Qy 1255 AGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1314  
Db 1435 ----- 1434  
Qy 1315 AGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1374  
Db 1435 -----ACTGATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTC 1471  
Qy 1375 CCAGGAAAACTCTTTTCATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAG 1434  
Db 1472 CTAGGAAAACTCTTTTCATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAG 1531  
Qy 1435 TGTTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAG 1494  
Db 1532 TGTTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAG 1591  
Qy 1495 AGCCCAAACTGGGAACTTCGAGGCGCATTCGAGTATTCGAGTATTCGAGTATTCGAG 1554  
Db 1552 AGCCCAAACTGGGAACTTCGAGGCGCATTCGAGTATTCGAGTATTCGAGTATTCGAG 1648  
Qy 1555 TGGATTCGATCAAAAGTCAAAAGAACTGATTCGATTCGATTCGATTCGATTCGATTC 1614  
Db 1649 TGGATTCGATCAAAAGTCAAAAGAACTGATTCGATTCGATTCGATTCGATTCGATTC 1708  
Qy 1615 CAGCAAAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1674  
Db 1709 CAGCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1768

Qy 1675 GAATGTTTTTGGCTCAGATTCCTCTTCTGAAACAGAGAAACCTAGGGGTATGAGCCA 1734  
Db 1769 GA-----GCCGGGCTTCATCTCTTCTGAAACAGAGAAACCTAGGGGTATGAGGTG 1822  
Qy 1735 CCATGGGGTTCATCAAAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1794  
Db 1823 ACATGGGGTTCATCAAAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1882  
Qy 1795 TTTAAACCAATCAGACAGAAAAACAAAAGGCTGTGTGAGATTCGATTCGAGGAG 1854  
Db 1883 TTTAAACCAATCAGACAGAAAAACAAAAGGCTGTGTGAGATTCGATTCGAGGAG 1942  
Qy 1855 GTGTGTGAGGAGCTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1914  
Db 1943 GTGTGTGAGGAGCTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2002  
Qy 1915 ATATCTAGTATGAAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1974  
Db 2003 ATATCTAGTATGAAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2062  
Qy 1975 GTTGTGATGACCACTCTCAGTATTCGATTCGATTCGATTCGATTCGATTCGATTC 2034  
Db 2063 GTTGTGATGACCACTCTCAGTATTCGATTCGATTCGATTCGATTCGATTCGATTC 2122  
Qy 2035 CCAGAAAAATCTGGGAAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2094  
Db 2123 CCAGAAAAATCTGGGAAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2182  
Qy 2095 AAATCTCCCAAAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2154  
Db 2183 AAATCTCCCAAAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2242  
Qy 2155 CTTGTGTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2214  
Db 2243 CTTGTGTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2302  
Qy 2215 TTTATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2274  
Db 2303 TTTATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2362  
Qy 2275 AATAGCTGAAAGAGAGATTAATTAATGATTCGATTCGATTCGATTCGATTCGATTC 2334  
Db 2363 AATAGCTGAAAGAGAGATTAATTAATGATTCGATTCGATTCGATTCGATTCGATTC 2422  
Qy 2335 TGTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2394  
Db 2423 TGTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2482  
Qy 2395 TTTCCCAATATCATGAGAAAGAACTGTGTGATTCGATTCGATTCGATTCGATTCGATTC 2454  
Db 2483 TTTCCCAATATCATGAGAAAGAACTGTGTGATTCGATTCGATTCGATTCGATTCGATTC 2542  
Qy 2455 CCTCTCTGTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2514  
Db 2543 CCTCTCTGTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2599  
Qy 2515 ATGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2574  
Db 2600 ATGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2659  
Qy 2575 GAAAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2634  
Db 2660 GAAAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2708  
Qy 2635 GATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2694  
Db 2709 -----GTTCCTTAATTCGACAGCTTTTGAATTCGATTCGATTCGATTCGATTC 2764  
Qy 2695 GCTACACGTTAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2754  
Db 2765 GCTACACGTTAATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 2824

Oy	2755	TTCGACGGAGGAGGTCTTCTATCAGTAATACTCCACCAAATGGCAACAACCTAGGAT	281.4
Db	2825	GTCACGGAGGAGATCTTCCATCATGTTACATATCACGAATGGTCAACAACACTAGGAT	288.4
Oy	2815	GGAGAAAATTGAAAAATTATCCAGACATCATCAAAGAATTACAGTGCTGCTTCCATC	287.4
Db	2885	GGAGAAAATTGAAAAATTATCCTGAATACATCAAAGAATTACAGTGCTTCTTCCATC	294.4
Oy	2875	CTTTGATGTTTTCTAATCCGACTCCATAATTTTCATGA	291.3
Db	2945	CTTCTGATGTTTTCTAATCCAATCTTAATTTTCAGTAA	298.3

RESULT 3  
US-08-834-108-3  
; Sequence 3, Application US/08834108

```

1 GENERAL INFORMATION:
2 APPLICANT: Dennis, James W
3 APPLICANT: Heffernan, Mike
4 APPLICANT: Fode, Carol
5 TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
6 NUMBER OF SEQUENCES: 14
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: BERSKIN & PARR
9 STREET: 40 King Street West
10 CITY: Toronto
11 STATE: Ontario
12 COUNTRY: Canada
13 ZIP: M5H 3Y2
14
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: Patentin Release #1.0, Version #1.30
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/834,108
22 FILING DATE:
23 CLASSIFICATION: 536
24 ATTORNEY/AGENT INFORMATION:
25 NAME: Kurdydyk, Linda M
26 REGISTRATION NUMBER: 34,971
27 REFERENCE/DOCKET NUMBER: 3153-210
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (416) 364-7311
30 TELEFAX: (416) 361-1398
31 INFORMATION FOR SEQ ID NO: 3:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 3447 base pairs
34 TYPE: nucleic acid
35 STRANDEDNESS: single
36 TOPOLOGY: linear
37 MOLECULE TYPE: cDNA
38 ORIGINAL SOURCE:
39 ORGANISM: Mus musculus
40 DEVELOPMENTAL STAGE: Lymphoid cDNA library
41 IMMEDIATE SOURCE:
42 LIBRARY: Murine Lymphoid
43 CLONE: WGA-resistant chop clones
44 FEATURE:
45 NAME/KEY: 5'UTR
46 LOCATION: 1..205
47 FEATURE:
48 NAME/KEY: CDS
49 LOCATION: 206..2980
50 FEATURE:
51 NAME/KEY: 3'UTR
52 LOCATION: 2981..3447
53
54 US-08-834-108-3
55
56 Query Match 64.5%; Score 1879; DB 2; Length 3447;
57 Best Local Similarity 80.9%; Pred.No.0;
58 Matches 2362; Conservative 0; Mismatches 410; Indels 147; Gaps 9;

```

Oy	ATGGGACACCTGCATCGGGGGAGAAATCGAGGATTTTAAAGTWGGAAATCTGCTTGTA	60
Db	ATGGCGGCGTGCACTGGGAGAGGATCGAGGACTTTAAAGTTGGAAATCTACTGGTA	265
Oy	GGATCACTTGGTGGTGTCTACAGAGCTGAGCCATTACACTGGTTTGGAAAGTTGCATC	120
Db	GGATCACTTGGTGGTGTCTACAGAGCTGAGCCATTACA	325
Oy	AAAAATGATAGTAAGAAAGCCATGTACAAAGACGAAATGCTACAGAGATCCAAATGAG	180
Db	AAAAATGATAGTAAGAAAGCCATGTACAAAGCTGAAATGCTACAGAGATCCAAATGAG	385
Oy	GTGAAATATCACTTGGCAATTGAAACATCTTCTATCTTGGAGCTTTATATATTTGAA	240
Db	GTGAAATATCACTTGGCAATTGAAACACCCCTCTGTGGAGCTCTATATATCTTTGAA	445
Oy	GATAGCAATTAATGTATCTGGTATTTGAAATGTGCCATATATGAGAAATGACAGGTAT	300
Db	GATAGCAATTAATGTATCTGGTATTTGAAATGTGCCATATGAGAAATGACAGGTAT	505
Oy	CTAAAGATATGAGTGAACCCCTTCTCAGAAATGAGCTCGACACTTCATGACACAGATC	360
Db	CTGAGAAACAGAAATGAGGCCCTTTCTCAGAAAGGAGCTGAGCACTTCATGACACAGAT	565
Oy	ATCACAGGGAATGTGTATCTTCATTCATGTGTATCTACACCGGGAACCTCACCTTTCT	420
Db	ATCACAGGAATGTGTATCTTCATTCATGTGTATCTACACCGGGAACCTCACCTCTCT	625
Oy	AACCTCTACTGACTCGGTATATATGAAATCAATCAAGATGCTGATTTTGGGCTGGCAACTCA	480
Db	AACATCTTACTTACGCGGAATATGAAATATGCTGACTTTTGGACTAGCAACGAG	685
Oy	CTGAAATATGCAACATGAAAGCACTATATATGTAAGTCTCTAACTACATTTCAACA	540
Db	TTGAAATATGCAACATGAAAGCACTATATATGTAAGTCTCTATATATTTCAACA	745
Oy	GAAATTTGCCACTCGAAGTGCACATGCGCTTGAAATCTGATGTTTGGTCCCTGGCTGTATG	600
Db	GAAATTTGCCACTCGAAGTGCACATGAGCACTTGAAATCTGATATTTGGTCAATGGGCTGTATG	805
Oy	TTTTATATCACTTATATCGGGAAGCAACCTTGGACATGACATGACAGTCAAGAACATTA	660
Db	TTTTATATCACTTATATCGGGAAGCAACCTTTTACATGACATGACATGACAGTCAAGT	865
Oy	AATAAAGTATGATGGGAGATTAAGAATGCGATCTTTTGTCAATAAGAGGCCAAGAC	720
Db	AATAAAGTATGATGGGAGATTAAGAATGCGATCTTTTGTCAATAAGAGGCCAAGAC	925
Oy	CTTATATCACAGTACTCTGTAGAAATCAGACAGATCGTTTAAGTCTGTCTTCAGTATG	780
Db	CTTATATCACAGTACTCTGTAGAAATCAGTCAATCGGTTAAAGTCTGTCTTCAGTATG	985
Oy	GACCATCTTTATGTGCCGAATTTCTTCACAAAAAGTAAAGTTTAAAGATCTGTGGA	840
Db	GACCATCTTTATGTGCCGAATTTCTTCACAAAAAGTAAAGCTTAAAGATCTGTGAG	1045
Oy	GACCTCAATGTATATGGGACATGCCAATTTCTATCGAATTAAGCTTCTTCAGTAC	900
Db	GACCTCAATGTATATGGGACATGCTACCTTTTCCAAACATTAACGCTCTTCTGTGATCC	1105
Oy	AGTATTAAGTGTATTTATTTGACAAAGAGAGACTTTTGTATGTGTACGCCATCCCAAT	960
Db	AGTATTAAGTGTATTTATTTGACAAAGAGAGACTTTTGTATGTGTACGCCATCCCAAT	1162
Oy	AAAATTTACTGATTTTCAAAAAATTAATAATTCAGATGACTT---TTCTTCAGGAGATGA	1219
Db	AAAATTTACTGATTTTCAAAAAATTAATAATTCAGATGACTT---TTCTTCAGGAGATGA	1274
Oy	AAAGTTTTTATCTCAGTGGGGAAT-----CAAGAAACAGTAAATATGAGAGGGGA	1074
Db	AAAGTTTTTATCTCAGTGGGGAATTTCCAGACATGATTAAGTATGAGGAGCGGGG	1279

QY 1075 AAGTAAATTCAGAGTCAGAGAAAGGCCAATCTTCGATACCTTCGTAGAGCTTATTC 1134  
 DB 1280 AAGAGATTTGAAGATTCAGAGAGAGCGGCATCTTCGATACCTTCGAGAGCTCATTC 1339  
 QY 1135 TCGTAAATTCAGAGCTTCTAATAGTCACTGCTCAAGCAAAAACATATCAATGAGACA 1194  
 DB 1340 TCGTAAAGCCAGCCCTCTAA--TCAGTCTCGAGCAAAAACATATCAAGAGAGCT 1396  
 QY 1195 TGTCACTCAGCAAAATGCTTTCAGTGTCCAAAAGATCAGAGAGAGTGAATAAGAG 1254  
 DB 1397 TGTCACTCAGTGAATATGCTTTCAGAGCCTTAGAAGATC----- 1434  
 QY 1255 AGGTACTCAACCAAGACAACTATCCAACTTTTAACTTCTTTAAAGAAAAGACATCC 1314  
 DB 1435 ----- 1434  
 QY 1315 AGTATGTCGATCTTTTGAAGACCTGATTAACATCAAGCACTTCGATCATCTTTGT 1374  
 DB 1435 -----ACTGATGAATAATCAACAGATTCGATCATCTTGT 1471  
 QY 1375 CCAGGAAAAATCTCTTTTCATTTGAGAGCCGACACTCAGACTGAAAACCGTACAACAG 1434  
 DB 1472 CTAGGAAAAATCTCTTTTCATTTGAGAGCCGACACTCAGATGGAATGATAGAGAG 1531  
 QY 1435 TGGTTTGGGAATCTGCAATTAATGCTCATTTAAGAAAACTAAGATATGACAGCATC 1494  
 DB 1532 TGGTTTGGGAATCTGCAATTAATGCTCATTTAAGAAAACTAAGATATGACAGCATC 1591  
 QY 1495 AGCCCAACCGGAGCTTCAGAGGCGCATTCAGATTTGAGAGAGCACTCAAAAATGCGC 1554  
 DB 1552 AGCCCAACAGAGATTTCCAGAGCTATCCAGATTTGC--AGAGACGTTTACGAAACGCT 1648  
 QY 1555 TGGAGTATACAAAAGTCAAAAAGAACTGTGATGCTTCTGATTAATGCACTTCTGAAA 1614  
 DB 1649 TGGAGTATACAAAAGTCAAAAAGAACTGTGATGCTTCTGATTAATGCACTTCTGAAA 1708  
 QY 1615 CAGCAAAATACATGAATATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAA 1674  
 DB 1709 CAGCTGAGTGCATGAATATATATGACTGCACTTCAAGTAAACCTGAGATTAATCCAA 1768  
 QY 1675 GAAATGTTTTTGGTCAATCTCTTTTGAACAGAGAGCACTAGGGGTATGAGCCA 1734  
 DB 1769 GA-----GCCGGGCTTACATCTCTCATTTGAAACAAAGAAATGAAGTATGAGTGG 1822  
 QY 1735 CCATGGGGTATCAGATGATGATTAAGAGCACTTACCTCCGTTGGTGTCCACAG 1794  
 DB 1823 AACCTGGGTATCAGAAACCTTACCTTAAGATATATCTCTCTGATGATGCTCA 1882  
 QY 1795 TTTAAACCAATCAGACAGAAAAACAAAAAGCTGTGTGAGCATCTTGAATTCAGAGAG 1854  
 DB 1883 TTTAAAGCCAAATCAGACAGAAAAACAAAAAGCTGTGTGAGCATCTTGAATTCAGAGAG 1942  
 QY 1855 GTGTGTGTGAGCTTGTAAAGAGATGATCTCAAGAAATATGTGAATAAGTCTTTCAG 1914  
 DB 1943 GTGTGTGTGAGCTTGTAAAGAGATGATCTCAAGAAATATGTGAATAAGTCTTTCAG 2002  
 QY 1915 ATATCTAGTGAATGAATATGATGATCTTATTTATTCAAATGTGTGTGAGTCTTCT 1974  
 DB 2003 ATATCTAGTGAATGAATATGATGATCTTATTTATTCAAATGTGTGTGAGTCTTCT 2062  
 QY 1975 CTGTCTGATAGACCACTCTACCTACTGACCAATCAGTATGATGATGATGATGATGAT 2034  
 DB 2063 CTGTCTGATAGACCTCTCTCTCTACTGACCAATCAGTATGATGATGATGATGATGAT 2122  
 QY 2035 CCAGAAAAATATCTGGGAAAAATATCAATATGCTTCCAGAGTTCAGATTCAGTATGAATCT 2094  
 DB 2123 CCAGAAAAATATCTGGGAAAAATATCAATATGCTTCCAGAGTTCAGATTCAGTATGAATCT 2182  
 QY 2095 AAATCTCCCAAAATCCTTATTTTACAAGATATGATTAATGATTTTGAATGAGATTC 2154  
 DB 2183 AAATCTCCCAAAATCCTTATTTTACAAGATATGATTAATGATTTTGAATGAGATTC 2242  
 QY 2155 CCTGTGTGATTTTGAAGTGTGTGTATGATGGGGTAAAAATACAAAAACAGAAAT 2214

DB 2243 CCTGTGTGATTTTGAAGTGTGTGTATGATGAGCCAAATATCATTAACCTGAAAAAT 2302  
 QY 2215 TTCAATTCAGTATTTGAAGAACAGAGAGTCTTCACTTTTAAAAAGTAAAGTGAAGT 2274  
 DB 2303 TTTAATTCACATATTTGAGAAAAACAGGATATCTTTAATTTTAAAAATATGAATGAGT 2362  
 QY 2275 AATAGCTTGAAGAGAGATTAATAATGATATGAGCAATGCTAATGAGGTCATGAT 2334  
 DB 2363 ACCAGCTTGAAGAGAGTAAAGATATATGAGCAATGCTAATGAGGTCATGAT 2422  
 QY 2335 TGTATGAGCTGTGATTCATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2394  
 DB 2423 TGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2482  
 QY 2395 TTCCCAATATCATATGAG 2454  
 DB 2483 TTCCCTATATGATGAG 2542  
 QY 2455 CCTCTTCTGTGATTCAAATTTACCAACGAGAGATGAGATCTTTTCAACAGAGATGTC 2514  
 DB 2543 CCTC---CTGTGAGCCCAAGCTGTGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2599  
 QY 2515 ATGATATGAGTGTCTTTTTCACACAGAGCACTTAAATCCCTTATGATGATGAT 2574  
 DB 2600 GTGAATATGAGTGTCTTTTTCACACAGAGCACTTAAATCCCTTATGATGATGAT 2659  
 QY 2575 GAAGAGCTGTGTCTTACCACTCAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2634  
 DB 2660 GAGAGATGAGTGTCTTACCACTCAGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 2708  
 QY 2635 GATGTCTTCTTAAATCAGACCACTTTTGAATCTTTTGTGAATAATGATGATGAT 2694  
 DB 2709 -----GTCTTCTTAAATCAGACAGCTTTTGAATCTTTTGTGAATAATGATGATGAT 2764  
 QY 2695 GCTTACACATTAATGATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2754  
 DB 2765 GCTTACACATTAATGATGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2824  
 QY 2755 GTGAGAGAGAGAGTGTCTTATCAGTTATACCTACCAATATGATGATGATGATGAT 2814  
 DB 2825 GTGAGAGAGAGAGTGTCTTATCAGTTATACCTACCAATATGATGATGATGATGAT 2884  
 QY 2815 GGAGAAAAATGAATAATTCAGACATTAATCAACAGAAATTAAGTGTGTGTGTGTGT 2874  
 DB 2885 GGAGAAAAATGAATAATTCAGACATTAATCAACAGAAATTAAGTGTGTGTGTGTGT 2944  
 QY 2875 CTTTGTATGTTTTCTAATCCAGCTCTTAATTTTCAATGA 2913  
 DB 2945 CTTGTATGTTTTCTAATCCAGCTCTTAATTTTCAGTAA 2983

RESULT 4  
 US-08-252-995D-5  
 ; Sequence 5, Application US/08252995D  
 ; Patent No. 5650501  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dennis, James W  
 ; APPLICANT: Heffernan, Mike  
 ; TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSES: BERSKIN & PARR  
 ; STREET: 40 King Street West  
 ; CITY: Toronto  
 ; STATE: Ontario  
 ; COUNTRY: Canada  
 ; ZIP: M5H 3Y2  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS



Qy	541	GAATTCACACTCGAATGCACTAGGCTTGAATCTGATGTTGGCTCGGCGTGATG	600
Db	746	GAATTTCMACTCGAATGCACTAGCACTTGAATCTGATATTGGCTCATTTGGCGTGATG	805
Qy	601	TTTATATACATTACTTAACTCGGAGACCAACCTTCGACACTGACAGTCGTAAGACACATTAA	660
Db	806	TCTTATACGTTACTTATTTGGAAGACCAACCTTTTGACACTGACAGTCGTAAGACACATTG	865
Qy	661	AATTAAGTAGTATTGGCAGATTATGAAATGCCATCTTTTGTGTCAATPAGAGCCCAAGAC	720
Db	866	AACAAAGTAGCTCCGGACAGATTATGAAATGCGACCCCTTTTGTGACGAGAGAGCCCAAGAC	925
Qy	721	CTTATTCACCGAGTACTTCGTGAGAAATCCAGCAATGTTTTAAGTCTGTCTTCAGTATTG	780
Db	926	CTTATTCACCGAGTACTTCGTGAGAAACCTTCGAGATGGTTAAGTCTGTCTTCGTGTGTG	985
Qy	781	GACCATCCCTTATATGTCGCCMAATTCCTTCAACAAAAGTAAAGATTAGAACCTGTGAA	840
Db	986	GACCATCCCTTATATGTCGCGMAATTCCTTCAACAAAAGTAAAGCGTAGGACCTGTAGAG	1045
Qy	841	GACTCAATTGATATGCGGCATGCCACAATTTCTTACTGCAATTACAGCTTCTTCCAGTACC	900
Db	1046	GACTCAATTGATATGCGGCATGCTACACTTTCACCAACAATTACAGCCTCTTCGTGTACC	1105
Qy	901	AGTATAGTGGTAGTTATTATTTGACAAAAGAAAGCTTTGATTTGGTCAGCCACTCCCAAT	960
Db	1106	AGTTTAGTGGCAAGCTTACTTGAC--AGAAAGCTTTTGTGTGTCAACCACTTCCAAAT	1162
Qy	961	AAATGACGTGATTTCCAAAGATAAAAGTTCAACGTATTTTCTTCAGAGATGGA	1020
Db	1163	AAATTTACTGTATTTCAAAATAAAATAATTTCAAGTACTT--TTCTTCAGAGATGGA	1219
Qy	1021	AACAGTTTATTACTCAGTGGGGAAT-----CAAGAAACCAAGTAATGTGAAAGGGA	1074
Db	1220	AGTATTTTGTGACTCAATGGGGAATTCAGAAACAAAGATTAATGATAGGGGACGGGGG	1279
Qy	1075	AGAGTATTTCAGATGCGAAGAAAGGCCACATCTCGATTAACCTTCGTAGAGCTTATTCC	1134
Db	1280	AGAGTGTATTGAAGATGCAAGAAAGAGGCCGATTTCTGATPACTCGCCAGAGCTCATTTCC	1339
Qy	1135	TCTGATAGATCTGGCACTTCTTAATAGTCAGTCTCAAGCAAAAACATATACATGGAAGCA	1194
Db	1340	TCTGATAGAGCCAGACCCCTCTTA--TCAGTCTGAGCAAAAACATACTCAGTAGAAAGT	1396
Qy	1195	TGTCACTCAGCAGAAATGCTTTCAGTGTCCAAAAGATCAGAGAGAGGTGAAATGAAAGAG	1254
Db	1397	TGTCACTCAGTAGAAATGCTTTCAGAAACCTTAGAAGATC--ACTGATGAAATATCAACAC	1453
Qy	1255	AGGTACTCACCACAGACAAACATGCCACATTTTAACTTTTAAAGAAAGACATCC	1314
Db	1454	AGGTATTCACCCACCAAAAGCAAGTGTCA--TGTTTTAACTTCATTTAAACCAAAACAGCC	1512
Qy	1315	AGTAGTCTGAGATCTTTGAAAGACGCTGATPACAATCAAGCACTCT	1360
Db	1513	AATAGTTAAGGATCTTTGAAAGACCGTATPATACTGAGCAGTAT	1558
RESULT 5			
US-08-834-108-5			
Sequence 5, Application US/08834108			
Patent No. 5976893			
GENERAL INFORMATION:			
APPLICANT: Dennis, James W			
APPLICANT: Heffernan, Mike			
APPLICANT: Fode, Carol			
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE			
NUMBER OF SEQUENCES: 14			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: BERESKIN & PARR			
STREET: 40 King Street West			
CITY: Toronto			
STATE: Ontario			
COUNTRY: Canada			

```

? ZIP: MSH 3Y2
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/834,108
? FILING DATE:
? CLASSIFICATION: 536
? ATTORNEY/AGENT INFORMATION:
? NAME: Kurdydk, Linda M
? REGISTRATION NUMBER: 34,971
? REFERENCE/DOCKET NUMBER: 3153-210
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (416) 364-7311
? TELEFAX: (416) 361-1398
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1600 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? ORIGINAL SOURCE:
? ORGANISM: Mus musculus
? DEVELOPMENTAL STAGE: Lymphoid cDNA library
? IMMEDIATE SOURCE:
? LIBRARY: Murine Lymphoid
? CLONE: WGA-resistant chop clones
? FEATURE:
? NAME/KEY: 5'UTR
? LOCATION: 1..205
? NAME/KEY: CDS
? LOCATION: 206..1597
? US-08-834-108-5

Query Match      33.3%; Score 970.4; DB 2; Length 1600;
Best Local Similarity 85.0%; Pred. No. 7e-263;
Matches 1161; Conservative 0; Mismatches 186; Indels 19; Gaps 6;

```

```

QY 481 CTGAAGATGCGACATGAAAGCACTTATGATGGAAGCTCTTACTTACTTACCA 540
DB 686 TTGAATATGCGACATGAAAGCACTTATGATGGAAGCTCTTACTTACTTACCA 745
QY 541 GAAATGCGACATGAAAGCACTTATGATGGAAGCTCTTACTTACTTACCA 600
DB 746 GAAATGCGACATGAAAGCACTTATGATGGAAGCTCTTACTTACTTACCA 805
QY 601 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 806 TCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 865
QY 661 AATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 866 AACAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 925
QY 721 CTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 926 CTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 985
QY 781 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 986 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1045
QY 841 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 1046 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1105
QY 901 AGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 1106 AGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1162
QY 961 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 1163 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1219
QY 1021 AACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1074
DB 1220 AACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1279
QY 1075 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1134
DB 1280 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1339
QY 1135 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194
DB 1340 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1396
QY 1195 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1254
DB 1397 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1453
QY 1255 AGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1314
DB 1454 AGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1512
QY 1315 AGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1360
DB 1513 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1558

RESULT 6
US-08-252-995D-1
; Sequence 1, Application US/08252995D
; Patent No. 5650501
; GENERAL INFORMATION:
; APPLICANT: Dennis, James W
; APPLICANT: Heffernan, Mike
; TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: BERESKIN & PARR  
 STREET: 40 King Street West  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5H 3Y2

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/252,995D  
 FILING DATE: 02-JUN-1994  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kurd/dyk, Linda M  
 REGISTRATION NUMBER: 34,971  
 REFERENCE/DOCKET NUMBER: 3153-96  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 364-7311  
 TELEFAX: (416) 361-1398

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1453 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 ORIGINAL SOURCE:  
 ORGANISM: Mus musculus  
 DEVELOPMENTAL STAGE: Lymphoid cDNA Library  
 IMMEDIATE SOURCE:  
 LIBRARY: Murine Lymphoid  
 CLONE: WGA-Resistant Chop clones  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 206..1453  
 NAME/KEY: 5'UTR  
 LOCATION: 1..205  
 US-08-252-995D-1

Query Match 31.7%; Score 924.2; DB 1; Length 1453;  
 Best Local Similarity 86.4%; Pred. No. 6,6e-250;  
 Matches 1071; Conservative 0; Mismatches 153; Indels 15; Gaps 4;

1 ATGCGACCTGCATCGGGAGAAAGATCGAGATTTTAAAGTTGGAATCTGCTGTAA 60  
 206 ATGCGCGCTGCATCGGGAGAGAGATCGAGACTTTAAGTTGGAATCTACTCGTAA 265  
 61 GATCATTTTCTGTGTCTTACAGAGCTGAGTCCATTCACTGTGTTGGAAGTTGCAATC 120  
 266 GATCATTTTCTGTGTCTTACAGAGCTGAGTCCATTCACTGTGTTGGAAGTTGCAATC 325  
 121 AAAATGATGATTAAGAAAGCAATGTAAGAAAGCAAGTGTAAAGAGTCCAAAATGAG 180  
 326 AAAATGATGATTAAGAAAGCAATGTAAGAAAGCAAGTGTAAAGAGTCCAAAATGAG 385  
 181 GTGAAATATCATTTGCCAATTTGAAACATCTTCTATCTTGGAGCTTTATTAATTTTGA 240  
 386 GTGAAATATCATTTGCCAATTTGAAACATCTTCTTGTGGAGCTTATTAATTTTGA 445  
 241 GATGCAATTTATGTGTATCTGTATTGAATGTGCCATTAATGAGAAATGAAACAGTAT 300  
 446 GATTAACAATTTATGTCTTACCTGTATTGAAATGTGCCAATTAATGAGAAATGAAACAGTAT 505  
 301 CTAAAGATTAAGTGAACCTCTTCAAGAAATGAAGCTGACACTTCAATGACCAAGATC 360  
 506 CTGAAGAAACGAATGAAGCTTTCTCAAGAAAGGAGCTAGGCACTTCAATGACCAAGAT 565  
 361 ATCACAGGATGTTGTATCTTCAATTTCTCATGTATTAATCAACCGGAGCTCAACTTCT 420

566 ATCACAGGATGTTGTATCTTCAATTTCTCATGTATTAATCAACCGGAGCTCAACTCTCT 625  
 421 AACCTCTTCTCATCTCTTAATTAAGAACATCAAGATTTGATTTGGGTGGCACTCA 480  
 626 AACATCTTACTTACCGGGAATTAAGAACATTAATTTGCTGACTTGGACTAGCAAGCAG 685  
 481 CTGAAATGCCACATGAAGAAGCACTATATATGTAAGTAACTCTTAATCAATTTGACCA 540  
 686 TTGAATATGCCACATGAAGAAGCACTATATATGTAAGTAACTCTTAATTAATTTGACCA 745  
 541 GAAATGCCACTGGAAGTGCACATGAGCTTGAATCTGATGTTTGTCTCTGAGCTGTATG 600  
 746 GAAATGCCACTGGAAGTGCACATGAGCTTGAATCTGATGTTTGTCTCTGAGCTGTATG 805  
 601 TTTTATACATTAATTAATGAGAGACACCTCTGACATGACACATGCAAGAAACATTA 660  
 806 TCTTATACATTAATTAATGAGAGACACCTCTTGAACATGACACATGCAAGAAACATTA 865  
 661 AATTAAGTATTAATGAGAGATTAATGAAATGCCATCTTTTGTCAATGAGGCCAAGAC 720  
 866 AACAAAGTATCTCTGACATTAATGAAATGCCATCTTTTGTCAAGAGGCCAAGAC 925  
 721 CTATATCACAGTACTTCTGTGAAGAAATCCAGACATGCTTTAACTGTCTTCAATTTG 780  
 926 CTATATCACAGTACTTCTGTGAAGAAATCCAGACATGCTTTAACTGTCTTCAATTTG 985  
 781 GACATCTTTTATATGTCGGAAATCTTCAAGAAATGAAGATTTAGAACTGAGAA 840  
 986 GACATCTTTTATATGTCGGAAATCTTCAAGAAATGAAGATTTAGAACTGAGAA 905  
 841 GACTCAATTAATTAATGAGAGATGACCAATTTTCTACTGCAATTAACCTTTCTCAATAC 960  
 1046 GACTCAATTAATTAATGAGAGATGACCAATTTTCTACTGCAATTAACCTTTCTCAATAC 1105  
 901 AGTAAAGTGTATGTTTATTTGAAGAAAGACCTTTGATGTTGACCACTCCCAAT 960  
 1106 AGTTAGTGTGAGGAGCTTACTTGAC--AGAGACTTTTGTGTGTCAACCACTTCCAAAT 1162  
 961 AAAATGATGATTTTCCAAAGAAATTAAGTCAAGATTTTCTTCTTCAAGAGATGGA 1020  
 1163 AAAATGATGATTTTCCAAAGAAATTAAGTCAAGATTTTCTTCTTCAAGAGATGGA 1219  
 1021 AACAGTTTATTAATCAAGTGGGGAAT-----CAAGAAACAGTAAATAGTGAAGGGA 1074  
 1220 AGTAAATTTTGTACTAATGAGGGAATCCAGAAAGAAAGCTTAATGAGGGAACGGGG 1279  
 1075 AGAGTAAATTAAGATGCAAGAAAGGAGGCAATTTCTGATACCTTGTGAGCTTAATTC 1134  
 1280 AGAGTAAATTAAGATGCAAGAAAGGAGGCAATTTCTGATACCTTGTGAGCTTAATTC 1339  
 1135 TCTGATTAATTAAGATGCAAGAAAGGAGGCAATTTCTGATACCTTGTGAGCTTAATTC 1194  
 1340 TCTGATTAATTAAGATGCAAGAAAGGAGGCAATTTCTGATACCTTGTGAGCTTAATTC 1396  
 1195 TCTCACTAGCAAGAAATGCTTCAAGTGTCCAAAGATCA 1233  
 1397 TCTCACTAGCAAGAAATGCTTCAAGGCTTAAGAAATCA 1435

RESULT 7  
 US-08-834-108-1  
 ; Sequence 1, Application US/08834108  
 ; Patent No. 5976893  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dennis, James W  
 ; APPLICANT: Helfferman, Mike  
 ; APPLICANT: Fode, Carol  
 ; TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: BERESKIN & PARR  
 ; STREET: 40 King Street West  
 ; CITY: Toronto

```

STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/834,108
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Kurdydyk, Linda M
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 3153-210
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
FAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1453 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Mus musculus
DEVELOPMENTAL STAGE: Lymphoid cDNA Library
LIBRARY: Murine Lymphoid
CLONE: MCA-Resistant Chop Clones
FEATURE:
NAME/KEY: CDS
LOCATION: 206..1453
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..205
US-08-834-108-1

```

Query Match 31.7%; Score 924.2; DB 2; Length 1453;  
Best Local Similarity 86.4%; Pred. No. 6.6e-250;  
Matches 1071; Conservative 0; Mismatches 153; Indels 15; Gaps 4;

```

QY 1 ATGGCGACCTGTCATCGGGAGAGATCGAGATTTTAAAGTTGAATCTGCTGGTAA 60
DB 206 ATGGCGGCGTCGATCGGGAGAGATCGAGATTTTAAAGTTGAATCTGCTGGTAA 265
QY 61 GATCACTTTGCTGCTGCTCAAGAGCTGATCTTCACTGCTTTGGAAGTTGCAATC 120
DB 266 GATCACTTTGCTGCTGCTCAAGAGCTGATCTTCACTGCTTTGGAAGTTGCAATC 325
QY 121 AAAATGATGATAGAGAAAGCCATGTAACAAGAGAGATGTAAGAGAGTCCAAATAG 180
DB 326 AAAATGATGATAGAGAAAGCCATGTAACAAGAGAGATGTAAGAGAGTCCAAATAG 385
QY 181 GTGAATAATCATTTGCCAATTTAAACAATCTTCTATCTTGAGACTTTATTAATTTTGA 240
DB 386 GTGAATAATCATTTGCCAATTTAAACAATCTTCTATCTTGAGACTTTATTAATTTTGA 445
QY 241 GATGCAATTTATGATCTGATCTGATTTAAGAAATGTCCTTAATGGAATGAAACAGTNT 300
DB 446 GATGCAATTTATGATCTGATCTGATTTAAGAAATGTCCTTAATGGAATGAAACAGTNT 505
QY 301 CTAAAGATAGAGTAAACCTTCTCAGAAATGAGCTCGACATTCATGACAGATC 360
DB 506 CTGAAGAACAGATAGAGCTTCTTCAAGAAAGAGAGAGCACTTATGACAGAT 565
QY 361 ATCAAGAGATTTGATTTTATTTCTCATGTAATGTAACACCGGAGCTCACTTTCT 420
DB 566 ATCAAGAGATTTGATTTTATTTCTCATGTAATGTAACACCGGAGCTCACTTTCT 625
QY 421 AACCTCTACTGACTGTAATGTAACATCAAGATTTGCTGATTTTGGGCTGCAACTCA 480

```

```

DB 626 AACATTTTACTTACGGGAAATGAAACATTAATAATTTGACTTGTGACTTACGACGACG 685
QY 481 CTGAAAATGCCAATGAAAAGCACTATACATTATGTAAGTCTTAATCACTATTTCCACCA 540
DB 686 TTGAATATGCCAATGAAAAGCACTATACCTGTGGGACCTCTAATATATTTCCACCA 745
QY 541 GAAATGGCACTGAAAGTGCACATGAGCTTGAATCTGAATGTTTGGTCCCTGGGCTGTAG 600
DB 746 GAAATGGCACTGAAAGTGCACATGAGCTTGAATCTGAATGTTTGGTCCCTGGGCTGTAG 805
QY 601 TTTATACATTACTTATCGGAGACACACCTTCGACACTGACACAGTCAAGAACATTA 660
DB 806 TCTATAGCTTACTTATGGAAGACCACTTTTGACACGACAGCTCAAGAACATTA 865
QY 661 AATTAAGTATGTTGGCAGATTTATGAATGCCATTTTGTCAATAGAGCCCAAGAC 720
DB 866 AACAAAGTATGCTGCGAGATTTATGAATGCCAGCTTTTGTCAAGAGCCCAAGAC 925
QY 721 CTATATCACAGTTCCTGTAAGAAATCCAGAGATCGTTAAGTCTGCTCAATTTG 780
DB 926 CTATATCACAGTTCCTGTAAGAAATCCAGAGATCGTTAAGTCTGCTCAATTTG 985
QY 781 GACCATCTTTATGCTCCGAAATTTCTCAACAAAAGTAAAGATTAGAACTGTGNA 840
DB 986 GACCATCTTTATGCTCAAGAAATCTTCAACAAAAGTAAAGATGTAAGCTGTAG 1045
QY 841 GACTCAATGATAGTGGGCAATCCCAATTTCTACTGCAATTAACGCTTTCTCCAGTACC 900
DB 1046 GACTCAATGATAGTGGGCAATCCCAATTTCTACTGCAATTAACGCTTTCTCCAGTACC 1105
QY 901 AGTATAGTGTATGTTATTTGACAAAAGAGACTTTGATTTGGCAGCACTCCCAAT 960
DB 1106 AGTATAGTGTATGTTATTTGACAAAAGAGACTTTGATTTGGCAGCACTCCCAAT 1162
QY 961 AAAATGACTGATTTTCCAAAGATTAAGTTCAACTGATTTTCTTCTCAGAGATGGA 1020
DB 1163 AAAATGACTGATTTTCCAAAGATTAAGTTCAACTGATTTTCTTCTCAGAGATGGA 1219
QY 1021 AACAGTTTATTAATCACTATGCGGGAAT-----CAAGAACAGTAATAGTGAAGGGA 1074
DB 1220 AGTAATTTTGTACTCAATAGGGAATTCAGAAACAAAGAGTAAATAGTAGGGAACGGGGG 1279
QY 1075 AGAGTAATTAAGATGCAAGAAAGGCAATTCATCTGATCTTCTGTAAGTATTC 1134
DB 1280 AGAGTAATTAAGATGCAAGAAAGGCAATTCATCTGATCTTCTGTAAGTATTC 1339
QY 1135 TCTGATGATCTGGCACTTCTAATAGTCAATCTCAAGCAAAACATATACATGGAACGA 1194
DB 1340 TCTGATGATGATCTGGCACTTCTAATAGTCAATCTCAAGCAAAACATATACATGGAACGT 1396
QY 1195 TGTCACTCAGAGAAATGCTTCAAGTCCAAAGATCA 1233
DB 1397 TGTCACTCAGAGAAATGCTTCAAGTCCAAAGATCA 1435

```

RESULT 8  
US-09-016-434-649  
Sequence 649; Application US/09016434  
Patent No. 650038  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304

```

1 OPERATING SYSTEM: DOS
2 SOFTWARE: FASTED for Windows Version 2.0
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER: US/08/878,989
5 FILING DATE:
6 CLASSIFICATION: 435
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER:
9 FILING DATE:
10 ATTORNEY/AGENT INFORMATION:
11 NAME: Billings, Lucy U J
12 REGISTRATION NUMBER: 36,749
13 REFERENCE/DOCKET NUMBER: PF-0321 US
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: 415-855-0555
16 TELEFAX: 415-845-4166
17 TELEX:
18 INFORMATION FOR SEQ ID NO: 8:
19 SEQUENCE CHARACTERISTICS:
20 LENGTH: 2770 base pairs
21 TYPE: nucleic acid
22 STRANDEDNESS: single
23 TOPOLOGY: linear
24 IMMEDIATE SOURCE:
25 LIBRARY: HUVENOB01
26 CLONE: 39043
27 ?
28 ?
29 ?
30 ?
31 ?
32 ?
33 ?
34 ?
35 ?
36 ?
37 ?
38 ?
39 ?
40 ?
41 ?
42 ?
43 ?
44 ?
45 ?
46 ?
47 ?
48 ?
49 ?
50 ?
51 ?
52 ?
53 ?
54 ?
55 ?
56 ?
57 ?
58 ?
59 ?
60 ?
61 ?
62 ?
63 ?
64 ?
65 ?
66 ?
67 ?
68 ?
69 ?
70 ?
71 ?
72 ?
73 ?
74 ?
75 ?
76 ?
77 ?
78 ?
79 ?
80 ?
81 ?
82 ?
83 ?
84 ?
85 ?
86 ?
87 ?
88 ?
89 ?
90 ?
91 ?
92 ?
93 ?
94 ?
95 ?
96 ?
97 ?
98 ?
99 ?
100 ?
101 ?
102 ?
103 ?
104 ?
105 ?
106 ?
107 ?
108 ?
109 ?
110 ?
111 ?
112 ?
113 ?
114 ?
115 ?
116 ?
117 ?
118 ?
119 ?
120 ?
121 ?
122 ?
123 ?
124 ?
125 ?
126 ?
127 ?
128 ?
129 ?
130 ?
131 ?
132 ?
133 ?
134 ?
135 ?
136 ?
137 ?
138 ?
139 ?
140 ?
141 ?
142 ?
143 ?
144 ?
145 ?
146 ?
147 ?
148 ?
149 ?
150 ?
151 ?
152 ?
153 ?
154 ?
155 ?
156 ?
157 ?
158 ?
159 ?
160 ?
161 ?
162 ?
163 ?
164 ?
165 ?
166 ?
167 ?
168 ?
169 ?
170 ?
171 ?
172 ?
173 ?
174 ?
175 ?
176 ?
177 ?
178 ?
179 ?
180 ?
181 ?
182 ?
183 ?
184 ?
185 ?
186 ?
187 ?
188 ?
189 ?
190 ?
191 ?
192 ?
193 ?
194 ?
195 ?
196 ?
197 ?
198 ?
199 ?
200 ?
201 ?
202 ?
203 ?
204 ?
205 ?
206 ?
207 ?
208 ?
209 ?
210 ?
211 ?
212 ?
213 ?
214 ?
215 ?
216 ?
217 ?
218 ?
219 ?
220 ?
221 ?
222 ?
223 ?
224 ?
225 ?
226 ?
227 ?
228 ?
229 ?
230 ?
231 ?
232 ?
233 ?
234 ?
235 ?
236 ?
237 ?
238 ?
239 ?
240 ?
241 ?
242 ?
243 ?
244 ?
245 ?
246 ?
247 ?
248 ?
249 ?
250 ?
251 ?
252 ?
253 ?
254 ?
255 ?
256 ?
257 ?
258 ?
259 ?
260 ?
261 ?
262 ?
263 ?
264 ?
265 ?
266 ?
267 ?
268 ?
269 ?
270 ?
271 ?
272 ?
273 ?
274 ?
275 ?
276 ?
277 ?
278 ?
279 ?
280 ?
281 ?
282 ?
283 ?
284 ?
285 ?
286 ?
287 ?
288 ?
289 ?
290 ?
291 ?
292 ?
293 ?
294 ?
295 ?
296 ?
297 ?
298 ?
299 ?
300 ?
301 ?
302 ?
303 ?
304 ?
305 ?
306 ?
307 ?
308 ?
309 ?
310 ?
311 ?
312 ?
313 ?
314 ?
315 ?
316 ?
317 ?
318 ?
319 ?
320 ?
321 ?
322 ?
323 ?
324 ?
325 ?
326 ?
327 ?
328 ?
329 ?
330 ?
331 ?
332 ?
333 ?
334 ?
335 ?
336 ?
337 ?
338 ?
339 ?
340 ?
341 ?
342 ?
343 ?
344 ?
345 ?
346 ?
347 ?
348 ?
349 ?
350 ?
351 ?
352 ?
353 ?
354 ?
355 ?
356 ?
357 ?
358 ?
359 ?
360 ?
361 ?
362 ?
363 ?
364 ?
365 ?
366 ?
367 ?
368 ?
369 ?
370 ?
371 ?
372 ?
373 ?
374 ?
375 ?
376 ?
377 ?
378 ?
379 ?
380 ?
381 ?
382 ?
383 ?
384 ?
385 ?
386 ?
387 ?
388 ?
389 ?
390 ?
391 ?
392 ?
393 ?
394 ?
395 ?
396 ?
397 ?
398 ?
399 ?
400 ?
401 ?
402 ?
403 ?
404 ?
405 ?
406 ?
407 ?
408 ?
409 ?
410 ?
411 ?
412 ?
413 ?
414 ?
415 ?
416 ?
417 ?
418 ?
419 ?
420 ?
421 ?
422 ?
423 ?
424 ?
425 ?
426 ?
427 ?
428 ?
429 ?
430 ?
431 ?
432 ?
433 ?
434 ?
435 ?
436 ?
437 ?
438 ?
439 ?
440 ?
441 ?
442 ?
443 ?
444 ?
445 ?
446 ?
447 ?
448 ?
449 ?
450 ?
451 ?
452 ?
453 ?
454 ?
455 ?
456 ?
457 ?
458 ?
459 ?
460 ?
461 ?
462 ?
463 ?
464 ?
465 ?
466 ?
467 ?
468 ?
469 ?
470 ?
471 ?
472 ?
473 ?
474 ?
475 ?
476 ?
477 ?
478 ?
479 ?
480 ?
481 ?
482 ?
483 ?
484 ?
485 ?
486 ?
487 ?
488 ?
489 ?
490 ?
491 ?
492 ?
493 ?
494 ?
495 ?
496 ?
497 ?
498 ?
499 ?
500 ?
501 ?
502 ?
503 ?
504 ?
505 ?
506 ?
507 ?
508 ?
509 ?
510 ?
511 ?
512 ?
513 ?
514 ?
515 ?
516 ?
517 ?
518 ?
519 ?
520 ?
521 ?
522 ?
523 ?
524 ?
525 ?
526 ?
527 ?
528 ?
529 ?
530 ?
531 ?
532 ?
533 ?
534 ?
535 ?
536 ?
537 ?
538 ?
539 ?
540 ?
541 ?
542 ?
543 ?
544 ?
545 ?
546 ?
547 ?
548 ?
549 ?
550 ?
551 ?
552 ?
553 ?
554 ?
555 ?
556 ?
557 ?
558 ?
559 ?
560 ?
561 ?
562 ?
563 ?
564 ?
565 ?
566 ?
567 ?
568 ?
569 ?
570 ?
571 ?
572 ?
573 ?
574 ?
575 ?
576 ?
577 ?
578 ?
579 ?
580 ?
581 ?
582 ?
583 ?
584 ?
585 ?
586 ?
587 ?
588 ?
589 ?
590 ?
591 ?
592 ?
593 ?
594 ?
595 ?
596 ?
597 ?
598 ?
599 ?
600 ?
601 ?
602 ?
603 ?
604 ?
605 ?
606 ?
607 ?
608 ?
609 ?
610 ?
611 ?
612 ?
613 ?
614 ?
615 ?
616 ?
617 ?
618 ?
619 ?
620 ?
621 ?
622 ?
623 ?
624 ?
625 ?
626 ?
627 ?
628 ?
629 ?
630 ?
631 ?
632 ?
633 ?
634 ?
635 ?
636 ?
637 ?
638 ?
639 ?
640 ?
641 ?
642 ?
643 ?
644 ?
645 ?
646 ?
647 ?
648 ?
649 ?
650 ?
651 ?
652 ?
653 ?
654 ?
655 ?
656 ?
657 ?
658 ?
659 ?
660 ?
661 ?
662 ?
663 ?
664 ?
665 ?
666 ?
667 ?
668 ?
669 ?
670 ?
671 ?
672 ?
673 ?
674 ?
675 ?
676 ?
677 ?
678 ?
679 ?
680 ?
681 ?
682 ?
683 ?
684 ?
685 ?
686 ?
687 ?
688 ?
689 ?
690 ?
691 ?
692 ?
693 ?
694 ?
695 ?
696 ?
697 ?
698 ?
699 ?
700 ?
701 ?
702 ?
703 ?
704 ?
705 ?
706 ?
707 ?
708 ?
709 ?
710 ?
711 ?
712 ?
713 ?
714 ?
715 ?
716 ?
717 ?
718 ?
719 ?
720 ?
721 ?
722 ?
723 ?
724 ?
725 ?
726 ?
727 ?
728 ?
729 ?
730 ?
731 ?
732 ?
733 ?
734 ?
735 ?
736 ?
737 ?
738 ?
739 ?
740 ?
741 ?
742 ?
743 ?
744 ?
745 ?
746 ?
747 ?
748 ?
749 ?
750 ?
751 ?
752 ?
753 ?
754 ?
755 ?
756 ?
757 ?
758 ?
759 ?
760 ?
761 ?
762 ?
763 ?
764 ?
765 ?
766 ?
767 ?
768 ?
769 ?
770 ?
771 ?
772 ?
773 ?
774 ?
775 ?
776 ?
777 ?
778 ?
779 ?
780 ?
781 ?
782 ?
783 ?
784 ?
785 ?
786 ?
787 ?
788 ?
789 ?
790 ?
791 ?
792 ?
793 ?
794 ?
795 ?
796 ?
797 ?
798 ?
799 ?
800 ?
801 ?
802 ?
803 ?
804 ?
805 ?
80
```

Oy	173	AAATGAGTGAAATAATCATTTGCCAATTGAAACCTCTCATCTTGGAGCTTATTA	232
Db	483	ACAAAGAAATAGAGCTTCAAGAAATCTCTCATATTAAGCATGTAGTGAGTTTAC	542
Oy	233	ATTTTGAAGTAGCAATTATGTGTATCTGGTATTAGAAATGCGCATTAATGAGAA	292
Db	543	ACTTCGAGGACAAAGAAACATTTTACATCTCTTGGAAATCTGACATGAAAGT	601
Oy	293	ACAGTATCTAAGAAATAGAGTGAACCCCTCTCAGAAATAGAACTGACATTTAC	352
Db	602	--GCTCATTTTGAAGAACGAAAGGTGTGACAGAGCCGAAAGTTGATCTAC	659
Oy	353	ACAGATCATCACAGGATGTTGTATCTTCATCTCATGTATACACCGGAGCTCA	412
Db	660	GGCAGATTGTGTGAGCTGAAATACCTTCATGAAACAAAGATCTTGACAGAT	719
Oy	413	CACTTCTAACCTCTCATGCTGTGTATATGAACATCAAGATTTGGAGCTG	472
Db	720	AACTTAGGAACTTTTATTTATTAATGACCATGAACTTAAAGTTGGGAGCTT	779
Oy	473	CAACTCACTGAAATGCCACATGAAAGACCTATATATGTGAACTCTCACTACA	532
Db	780	CAGCCAGGCTAAGAACCTTGGAAACAGAAAGAAACGATGTGTGTATGCCCA	839
Oy	533	TTTTCACAGAAATTGCCATCGAGTGCACATGGCCCTGATCTGATGTTGGTCC	592
Db	840	TCTCTCTGAAAGTCTTCAACAAACAAAGACATGGCTGTGAATCAGACATTT	899
Oy	593	GCTGTATGTTTATACATTACTTATCGGAGAACACCTCTGCAGCACTGACAC	652
Db	900	GCCTGTATATGATATACAAATGTATATAGGAGGCCCATTTGAAACTCAAA	959
Oy	653	ACACATTAATAAAGTATGTATGGCAGATTAATGAATGCCATCTTTTGTCCA	712
Db	960	AAACTTTATAGTGATAGGAAACAAAGATACAAATGCCCTCTCATTTGGCT	1019
Oy	713	CCAAGACCTTATTCACCAAGTATCTTGATGAATTCAGACAGATCGTTAATG	772
Db	1020	CCAAAGCATTAATTTGCTAGTATGTGTTCCAAACCCAGAGATGTGTCCAG	1079
Oy	773	CAGTATGACCATCTTTTATGT	796

DB 1080 ACATCATTCGACATGACTTTT 1103

## RESULT 10

US-09-272-796-8  
Sequence 8, Application US/09272796  
Patent No. 6207148  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Corley, Neil C.  
APPLICANT: Guebler, Karl G.  
APPLICANT: Lal, Preeti  
APPLICANT: Goli, Surya K.  
APPLICANT: Shah, Puri  
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN  
TITLE OF INVENTION: KINASES  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/272,796  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/878,989  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0321 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2770 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: HUVEHOB01  
CLONE: 39043  
US-09-272-796-8

Query Match 4.6%; Score 134.4; DB 3; Length 2770;  
Best Local Similarity 52.1%; Pred. No. 1.7e-27;  
Matches 325; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

DB 173 AAAATGAGTGAATAATACATTCGCAATGAACATCCCTTATCTTGGAGCTTTATACT 232  
DB 483 ACAAGAATATGAGCTTACAGAAATCTTCAATCAATGATGAGTTCAGACT 542  
DB 233 ATTTGAGATGACATTTATGTGTGATTAAGAAATGTCATATGAGAAATGA 292  
DB 543 ACTTGAAGAGAAAGAAACATTTACATTTCTTGGAATACAGAGTGAAGGTCAAG- 601  
DB 293 ACAGGTACTAAAGATAGAGGAACCTTCTCGAAGAAATGAGAGCTGACACTGATGC 352  
DB 602 --GCTCATATTTTGAAGCAAGAAAGGTGTGACAGACCAAGATTGATCTACTCA 659

DB 353 AACGATCATCAGAGGATGTGATCTTCTCATGTATACATCACCGGAGCTCA 412  
DB 660 GGCAGATGTGTCTGACATAAATACCTTCAATGACAGAAATCTTGACAGAGATCTCA 719  
DB 413 CACTTCTAACTCTCTACTGACTGTATATGAAACATCAAGATTTGCTGATTTGGGCTGG 472  
DB 720 AACTAGGGAATCTTTTATTAATGAAACCATGAAATGAAAGTGGGACCTTGCTGG 779  
DB 473 CAATCACTGAAATATCCACATGAAAGCATATATGATATGGAATCTTAATACATCA 532  
DB 780 CAGCCAGCTAGAACCTTGGAAACAGAAAGAGAAAGATATGTGTATCCCAATATATC 839  
DB 533 TTTCACAGAAATATGCACTGCAATGACATGAGCTTGAATCTGATGTGTGCTCTGG 592  
DB 840 TCTCTCCGAAAGTCTTCAACAAAGACATAGCTGTGATCAGACATTTGGGCTCTGG 899  
DB 593 GCTGTATGTTTATATCAATTAATCTTATGCGAGACCACTTCTGACATGACATGACAGA 652  
DB 900 GCTGTATGTTTATATCAATGTATCTAGGAGGCCCCCATTTGAAACTACAAATCTCAAG 959  
DB 653 ACACCTTAATTAAGTGTATGTCAGATATGAAATGCCATCTTTTGTCAATAGAG 712  
DB 960 AACTTATAGTGTATGAGGAGAGCAAGTATACAAATGCCCTCATGTGCTGCTCTG 1019  
DB 713 CCAAGACCTTATTCACCAAGTATCTTGTAGAAATCCAGACATGCTTAAATCTCTCTT 772  
DB 1020 CCAAGCACTTATATGCTAGTATGTGTCCAAAAACCCAGAGATGTCCTCCAGTTGATG 1079  
DB 773 CAGTATGACCATCTTTTATGT 796  
DB 1080 ACATCATTCGACATGACTTTT 1103

## RESULT 11

US-09-136-282-3  
Sequence 3, Application US/09136282  
Patent No. 6063609  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, KAREN  
APPLICANT: JACKSON, JEFFREY  
APPLICANT: HANSBURY, MICHAEL  
APPLICANT: NERURKAR, SANDHYA  
APPLICANT: KOSHAH, AMY  
APPLICANT: BOUZYK, MARK  
TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ratner & Prestia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,282  
FILING DATE: 20-AUG-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/056,112  
FILING DATE: 20-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-70231  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0700  
TELEX: 846169

; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2789 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; US-09-136-282-3

Query Match 4.6%; Score 134.4; DB 3; Length 2789;  
 Best Local Similarity 52.1%; Pred. No. 1.7e-27;  
 Matches 325; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

```

QY 173 AAAATAGGTGAAATATACATTCGCAATGAAATCCTCTCTATCTTGAGCTTTAACT 232
DB 500 ACAAGAAATAGAGCTTCACAGAAATCTTCATCATAGCATGATGAGCTTTACCACT 559
QY 233 ATTTGAGATGACATTTATGTGTATCTGATTTAGAAATGCGCATTAATGAGAAATGA 292
DB 560 ACTTGAGAGACAAAGAAACATTTACATCTCTTGGAATATGCAATGAGAGTCAATG- 618
QY 293 ACAAGTATCTAAAGATAGAGTGAACCTTCTCAGAAATGAGCTCGACATTCATGC 352
DB 619 --GCTCATATTTTGAAGAGAAAGAGTGTGACAGAGCCAGAAAGTTCATCTACCTCA 676
QY 353 ACCAGATCATCAGAGGATGTTGTATCTTCATCTCATGATGATACACACCGGAGCTCA 412
DB 677 GGCAGATGTGTCTGAGACTGAATACCTTCATGAACAGAAATCTTGACAGAGATCTCA 736
QY 413 CACTTCTAACCTCCTACTGACTCTGTAAATGAAACATCAAGATTGCTGATTTGGGCTGG 472
DB 737 AACTAGGGAACCTTTTATTAATGAAGCAGATGAACTAAAGTTGGGAGCTTGGCTGG 796
QY 473 CAATCACTGAAATGCCACATGAAAGACATATACATTAATGTGAACCTCTAATCA 532
DB 797 CAGCCAGGCTAGAACCTTTGAAACACAGAGAGAAAGATGTGTATGTAACCCCAAAATATC 856
QY 533 TTTCACAGAAATGCGCATCGAGAGTGCATGCGCTTGAATGTAATGTTGGTCCCTGG 592
DB 857 TCTCTCTGAAGTCTCAACAAAGAGATGCGTGTATGATGACATTTGGGCTGG 916
QY 593 GCTGTATGTTTATACATTAATCTTATCGGAGACCACTTCTGACACTGACAGTCAAGA 652
DB 917 GCTGTATGTTATACATTAATCTTATCGGAGAGGCCCATTTGAAATCAAAATCTCAAG 976
QY 653 ACAATTAATAAAGATGATGATGAGCAATTAAGAAATGCGATCTTTTGTCAATAGAG 712
DB 977 AAATCTATAGTGATAGAGGAAAGCAAGGATATACATGCGTCTCATTTGCTGCTCG 1036
QY 713 CCAAGACCTTATATCCAGCTTACTCTGTAGAAATCCAGACAGATCGTTTAAGCTGCTT 772
DB 1037 CCAAGACCTTAATGCTAGATGTTGTCCAAAAACCAAGAGATGCTCCAGTTTGAGATG 1096
QY 773 CAGTATGAGCAATCCTTTTATGT 796
DB 1097 ACATCATTCGACATGACTTTT 1120
  
```

RESULT 12  
 US-09-505-744-3  
 ; Sequence 3, Application US/09505744  
 ; Patent No. 6245544  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Karen M. Anderson  
 ; APPLICANT: Mark M. Bouzyk  
 ; APPLICANT: Michael J. Hansbury  
 ; APPLICANT: Jeffrey R. Jackson  
 ; APPLICANT: Sandhya S. Nerurkar  
 ; APPLICANT: Amy K. Roehak  
 ; TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)  
 ; FILE REFERENCE: GH-70231-D1  
 ; CURRENT APPLICATION NUMBER: US/09/505, 744  
 ; CURRENT FILING DATE: 2000-02-16

; EARLIER APPLICATION NUMBER: 09/136,282  
 ; EARLIER FILING DATE: 1998-08-20  
 ; EARLIER APPLICATION NUMBER: 60/056,112  
 ; EARLIER FILING DATE: 1997-08-20  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 2789  
 ; TYPE: DNA  
 ; ORGANISM: HOMO SAPIENS  
 ; US-09-505-744-3

Query Match 4.6%; Score 134.4; DB 3; Length 2789;  
 Best Local Similarity 52.1%; Pred. No. 1.7e-27;  
 Matches 325; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

```

QY 173 AAAATAGGTGAAATATACATTCGCAATGAAATCCTCTCTATCTTGAGCTTTAACT 232
DB 500 ACAAGAAATAGAGCTTCACAGAAATCTTCATCATAGCATGATGAGCTTTACCACT 559
QY 233 ATTTGAGATGACATTTATGTGTATCTGATTTAGAAATGCGCATTAATGAGAAATGA 292
DB 560 ACTTGAGAGACAAAGAAACATTTACATCTCTTGGAATATGCAATGAGAGTCAATG- 618
QY 293 ACAAGTATCTAAAGATAGAGTGAACCTTCTCAGAAATGAGCTCGACATTCATGC 352
DB 619 --GCTCATATTTTGAAGAGAAAGAGTGTGACAGAGCCAGAAAGTTCATCTACCTCA 676
QY 353 ACCAGATCATCAGAGGATGTTGTATCTTCATCTCATGATGATACACACCGGAGCTCA 412
DB 677 GGCAGATGTGTCTGAGACTGAATACCTTCATGAACAGAAATCTTGACAGAGATCTCA 736
QY 413 CACTTCTAACCTCCTACTGACTCTGTAAATGAAACATCAAGATTGCTGATTTGGGCTGG 472
DB 737 AACTAGGGAACCTTTTATTAATGAAGCAGATGAACTAAAGTTGGGAGCTTGGCTGG 796
QY 473 CAATCACTGAAATGCCACATGAAAGACATATACATTAATGTGAACCTCTAATCA 532
DB 797 CAGCCAGGCTAGAACCTTTGAAACACAGAGAGAAAGATGTGTATGTAACCCCAAAATATC 856
QY 533 TTTCACAGAAATGCGCATCGAGAGTGCATGCGCTTGAATGTAATGTTGGTCCCTGG 592
DB 857 TCTCTCTGAAGTCTCAACAAAGAGATGCGTGTATGATGACATTTGGGCTGG 916
QY 593 GCTGTATGTTTATACATTAATCTTATCGGAGACCACTTCTGACACTGACAGTCAAGA 652
DB 917 GCTGTATGTTATACATTAATCTTATCGGAGAGGCCCATTTGAAATCAAAATCTCAAG 976
QY 653 ACAATTAATAAAGATGATGATGAGCAATTAAGAAATGCGATCTTTTGTCAATAGAG 712
DB 977 AAATCTATAGTGATAGAGGAAAGCAAGGATATACATGCGTCTCATTTGCTGCTCG 1036
QY 713 CCAAGACCTTATATCCAGCTTACTCTGTAGAAATCCAGACAGATCGTTTAAGCTGCTT 772
DB 1037 CCAAGACCTTAATGCTAGATGTTGTCCAAAAACCAAGAGATGCTCCAGTTTGAGATG 1096
QY 773 CAGTATGAGCAATCCTTTTATGT 796
DB 1097 ACATCATTCGACATGACTTTT 1120
  
```

RESULT 13  
 US-09-136-282-1  
 ; Sequence 1, Application US/09136282  
 ; Patent No. 6063609  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANDERSON, KAREN  
 ; APPLICANT: JACKSON, JEFFREY  
 ; APPLICANT: HANSBURY, MICHAEL  
 ; APPLICANT: NERURKAR, SANDHYA  
 ; APPLICANT: ROEHAK, AMY  
 ; APPLICANT: BOUZYK, MARK  
 ; TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)

NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ratner & Prestlia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/136,282  
FILING DATE: 20-AUG-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/056,112  
FILING DATE: 20-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestlia, Paul F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-70231  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0700  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2783 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-136-282-1

Query Match 4.5%; Score 132.4; DB 3; Length 2783;  
Best Local Similarity 51.9%; Pred. No. 6.3e-27;  
Matches 324; Conservative 0; Mismatches 296; Indels 4; Gaps 1;  
QY 173 AAAATGAGTGAAATATCATTTGCCAATTGAAACCTCTTATCTTGAGCTTTATTAAT 232  
DB 500 ACAAGAAATAGAGCTTCACAGAAATCTTCATCATTAAGCATGATGACATTTTACCACT 559  
QY 233 ATTTGAGATGACATTAATGATGATCTGGTATTGAAAGTCCATTAATGAGAAATGA 292  
DB 560 ACTTCGAGGACAAAGAAACATTTACATCTCTTGAAATCTGAGATGAAAGCAATGCG 619  
QY 293 ACAGGTATCTAAAGATAGAGTGAACCTTCTCAGAAATGAAGCTGACACTTCATGCG 352  
DB 620 TCAATATTTGAAG---CAAGAAAGGTGTGACAGAGCCAGAAAGTTGATGATCACTCA 675  
QY 353 ACCAGATCATCAGAGGATGTTGATCTTCATCTCATGATTAATCAACCGGAGACTCA 412  
DB 676 GGAGATGTTGTGACATGAATACCTTCATGAACAAAGAAATCTTCCACAGAAATCTCA 735  
QY 413 CACTTTCTAACCTCTACTGCTGTAATATGAACATCAAGATGCTGATTTGGGCTGG 472  
DB 736 AACTAGGGAACCTTTTATTAATGAAGCATGAACATTAAGTTGGGGACTTGGGTCTGG 795  
QY 473 CAATCAACTGAAATATGCCATGAAAGCACTATATATGGAACCTCTTAATCA 532  
DB 796 CAGCAGGCTGAAGACCTTGAACACAGAAAGAAAGATATGATGATACCCCAATATATC 855  
QY 533 TTTCACAGAAATGCGCATGGAAGTGCATGCGCTTGAATCTGAATGTTGGTCTGG 592  
DB 856 TCTCTCCTGAAGTCTCTCAACAAACAAAGACATGCTGATGATCAAGCAATTTGGGCTGG 915  
QY 593 GCTGATGTTTATATCATTAATCTTAATCGGGAACCACTTGAACATGACACAGTCAAGA 652  
DB 916 GCTGTATATGATATCAATATGTTACTAGGAGGCCCCCAATTGAAACTACAAATCTCAAG 975

QY 653 ACACATTAATTAAGTAGATATGCGAGATTAATGAATGCCATCTTTTGTCAATAGAG 712  
DB 976 AACTTATAGTGATTAAGGAAGCAAGATTAATACAAAGCCGTCTCATCTGGCTCTG 1035  
QY 713 CCAAGACCTTATTCACAGATTAATGTTGAAATCCAGAGATCTTTAAGTCTGCT 772  
DB 1036 CCAAGACCTTATTAATGATGATGTTGTCAAAACCAAGAGATAGGCTTATGATG 1095  
QY 773 CAGTATTTGACATCTTTTATGT 796  
DB 1096 ACATCATTCAGACATGACTTTT 1119

RESULT 14  
US-09-505-744-1  
Sequence 1, Application US/09505744  
Patent No. 6245544  
GENERAL INFORMATION:  
APPLICANT: Karen M. Anderson  
APPLICANT: Mark M. Bouzyk  
APPLICANT: Michael J. Hanbury  
APPLICANT: Jeffrey R. Jackson  
APPLICANT: Sandhya S. Neturkar  
APPLICANT: Amy K. Roshak  
TITLE OF INVENTION: HUMAN SERUM INDUCIBLE KINASE (SNK)  
FILE REFERENCE: GH-70231-D1  
CURRENT APPLICATION NUMBER: US/09/505,744  
CURRENT FILING DATE: 2000-02-16  
EARLIER APPLICATION NUMBER: 09/136,282  
EARLIER FILING DATE: 1998-08-20  
EARLIER APPLICATION NUMBER: 60/056,112  
EARLIER FILING DATE: 1997-08-20  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 2783  
TYPE: DNA  
ORGANISM: HOMO SAPIENS  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (2720) (2721)  
US-09-505-744-1

Query Match 4.5%; Score 132.4; DB 3; Length 2783;  
Best Local Similarity 51.9%; Pred. No. 6.3e-27;  
Matches 324; Conservative 0; Mismatches 296; Indels 4; Gaps 1;  
QY 173 AAAATGAGTGAAATATCATTTGCCAATTGAAACCTCTTATCTTGAGCTTTATTAAT 232  
DB 500 ACAAGAAATAGAGCTTCACAGAAATCTTCATCATTAAGCATGATGACATTTTACCACT 559  
QY 233 ATTTGAGATGACATTAATGATGATCTGGTATTGAAAGTCCATTAATGAGAAATGA 292  
DB 560 ACTTCGAGGACAAAGAAACATTTACATCTCTTGAAATCTGAGATGAAAGCAATGCG 619  
QY 293 ACAGGTATCTAAAGATAGAGTGAACCTTCTCAGAAATGAAGCTGACACTTCATGCG 352  
DB 620 TCAATATTTGAAG---CAAGAAAGGTGTGACAGAGCCAGAAAGTTGATGATCACTCA 675  
QY 413 CACTTTCTAACCTCTACTGCTGTAATATGAACATCAAGATGCTGATTTGGGCTGG 472  
DB 736 AACTAGGGAACCTTTTATTAATGAAGCATGAACATTAAGTTGGGGACTTGGGTCTGG 795  
QY 473 CAATCAACTGAAATATGCCATGAAAGCACTATATATGGAACCTCTTAATCA 532  
DB 796 CAGCAGGCTGAAGACCTTGAACACAGAAAGAAAGATATGATGATACCCCAATATATC 855  
QY 533 TTTCACAGAAATGCGCATGGAAGTGCATGCGCTTGAATCTGAATGTTGGTCTGG 592





**This Page Blank (uspto)**